PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/19, C07K 14/52, 16/24, A61K 38/19

A1

(11) International Publication Number: WO 95/18858

(43) International Publication Date: 13 July 1995 (13.07.95)

(21) International Application Number: PCT/US94/14553

(22) International Filing Date: 28 December 1994 (28.12.94)

(30) Priority Data:

08/176,553	3 January 1994 (03.01.94)	US
08/185,607	21 January 1994 (21.01.94)	US
08/196,689	15 February 1994 (15.02.94)	US
08/223,263	4 April 1994 (04.04.94)	US
08/249,376	25 May 1994 (25.05.94)	US
08/348,657	2 December 1994 (02.12.94)	US
08/348,658	2 December 1994 (02.12.94)	US

(60) Parent Applications or Grants
(63) Related by Continuation

US	08/249,376 (CIP)
Filed on	25 May 1994 (25.05.94)
US	08/223,263 (CIP)
Filed on	4 April 1994 (04.04.94)
US	08/196,689 (CIP)
Filed on	15 February 1994 (15.02.94)
US	08/348,658 (CIP)
Filed on	2 December 1994 (02.12.94)
US	08/185,607 (CIP)
Filed on	21 January 1994 (21.01.94)
US	08/348,657 (CIP)
Filed on	2 December 1994 (02.12.94)
US	08/176,553 (CIP)
Filed on	3 January 1994 (03.01.94)

(71) Applicant *(for all designated States except US)*: GENENTECH, INC. [US/US]; 460 Point San Bruno Boulevard, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: EATON, Dan, L. [US/US]; 75 Knight Drive, San Rafael, CA 94901 (US). DE SAUVAGE, Frederic, J. [BE/US]; 166 Beach Park Boulevard, Foster City, CA 94404 (US).

(74) Agents: WINTER, Daryl, B. et al.; Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080-4990 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: THROMBOPOIETIN

(57) Abstract

Isolated thrombopoietin (TPO), isolated DNA encoding TPO, and recombinant or synthetic methods of preparing and purifying TPO are disclosed. Various forms of TPO are shown to influence the replication, differentiation or maturation of blood cells, especially megakaryocytes and megakaryocyte progenitor cells. Accordingly, these compounds may be used for treatment of thrombocytopenia.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

```
  1 tcttcctacccatctgctccccagagggctgcctgctgtgcacttgggtcctggagcccttctccacccggatagattcctcaccccttggcccgcctttg

101 ccccacccctactctgcccagaagtgcaagagcctaagccgcctccatggccccaggaaggattcaggggagaggccccaaacagggagccacgccagcca

                    -20                       -10
                 MetGluLeuThrGluLeuLeuLeuValValMetLeuLeuLeuThrAlaArgLeuThrLeuSerSerProAlaProProAlaCysAsp
201 gacaccccggccagaATGGAGCTGACTGAATTGCTCCTCGTGGTCATGCTTCTCCTAACTGCAAGGCTAACGCTGTCCAGCCCGGCTCCTCCTGCTTGTG

     10                        20                        30                          40
    LeuArgValLeuSerLysLeuLeuArgAspSerHisValLeuHisSerArgLeuSerGlnCysProGluValHisProLeuProThrProValLeuLeu
301 ACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCCCATGTCCTTCACAGCAGACTGAGCCAGTGCCCAGAGGTTCACCCTTTGCCTACACCTGTCCTGCT

                  50                        60                        70
    ProAlaValAspPheSerLeuGlyGluTrpLysThrGlnMetGluGluThrLysAlaGlnAspIleLeuGlyAlaValThrLeuLeuLeuGluGlyVal
401 GCCTGCTGTGGACTTTAGCTTGGGAGAATGGAAAACCCAGATGGAGGAGACCAAGGCACAGGACATTCTGGGAGCAGTGACCCTTCTGCTGGAGGGAGTG

              80                        90                        100
    MetAlaAlaArgGlyGlnLeuGlyProThrCysLeuSerSerLeuLeuGlyGlnLeuSerGlyGlnValArgLeuLeuLeuGlyAlaLeuGlnSerLeuLeu
501 ATGGCAGCACGGGGACAACTGGGACCCACTTGCCTCTCATCCTCCTGGGGCAGCTTTCTGGACAGGTCCGTCTCCTCCTTGGGGCCCTGCAGAGCCTCC

     110                       120                       130                        140
    GlyThrGlnLeuProProGlnGlyArgThrThrAlaHisLysAspProAsnAlaIlePheLeuSerPheGlnHisLeuLeuArgGlyLysValArgPhe
601 TTGGAACCCAGCTTCCTCCACAGGGCAGGACCACAGCTCACAAGGATCCCAATGCCATCTTCCTGAGCTTCCAACACCTGCTCCGAGGAAAGGTGCGTTT

                  150                       160                       170
    LeuMetLeuValGlyGlySerThrLeuCysValArgArgAlaProProThrThrAlaValProSerArgThrSerLeuValLeuThrLeuAsnGluLeu
701 CCTGATGCTTGTAGGAGGGTCCACCCTCTGCGTCAGGCGGGCCCCACCCACCACAGCTGTCCCCAGCAGAACCTCTCTAGTCCTCACACTGAACGAGCTC

                  180                       190                       200
    Pro[AsnArgThr]SerGlyLeuLeuGluThr[AsnPheThr]AlaSerAlaArgThrThrGlySerGlyLeuLeuLysTrpGlnGlnGlyPheArgAlaLysIle
801 CCAAACAGGACTTCTGGATTGTTGGAGACAAACTTCACTGCCTCAGCCAGAACTACTGGCTCTGGGCTTCTGAAGTGGCAGCAGGGATTCAGAGCCAAGA
```

FIG. 1A


210 220 230 240
ProGlyLeuLeuAsnGlnThrSerArgSerLeuAspGlnIleProGlyTyrLeuAsnArgIleHisGluLeuLeuAsnGlyThrArgGlyLeuPhePro
901 TTCCTGGTCTGCTGAACCAAACCTCCAGGTCCCTGGACCAAATCCCCGGATACCTGAACAGGATACACGAACTCTTGAATGGAACTCGTGGACTCTTTCC
250 260 270
GlyProSerArgArgThrLeuGlyAlaProAspIleSerSerGlyThrSerAspThrGlySerLeuProProAsnLeuGlnProGlyTyrSerProSer
1001 TGGACCCTCACGCAGGACCCTAGGAGCCCCGGACATTTCCTCAGGAACATCAGACACAGGCTCCCTGCCACCCAACCTCCAGCCTGGATATTCTCCTTCC
280 290 300
ProThrHisProProThrGlyGlnTyrThrLeuPheProLeuProProThrLeuProThrProValValGlnLeuHisProLeuLeuProAspProSerAla
1101 CCAACCCATCCTCCTACTGGACAGTATACGCTCTTCCCTCTTCCACCCACCTTGCCCACCCCTGTGGTCCAGCTCCACCCCCTGCTTCCTGACCCTTCTG
310 320 330
ProThrProThrProThrSerProLeuLeuAsnThrSerTyrThrHisSerGlnAsnLeuSerGlnGluGly
1201 CTCCAACGCCCACCCCTACCAGCCCTCTTCTAAACACATCCTACACCCACTCCCAGAATCTGTCTCAGGAAGGGTAAggttctcagacactgccgacatc
1301 agcattgtctcatgtacagctcccttccctgcagggcgcccctgggagacaactggacaagatttcctactttctcctgaaacccaaagccctggtaaaa
1401 gggatacacaggactgaaaagggaatcatttttcactgtacattataaaccttcagaagctatttttttaagctatcagcaatactcatcagagcagcta
1501 gctctttggtctattttctgcagaaatttgcaactcactgattctctacatgctctttttctgtgataactctgcaaaggcctgggctggcctggcagtt
1601 gaacagagggagagactaaccttgagtcagaaaacagagaaagggtaatttcctttgcttcaaattcaaggccttccaacgcccccatcccctttactat
1701 cattctcagtgggactctgatcccatattcttaacagatctttactcttgagaaatgaataagctttctctcagaaaaaaaaaaaaaaaaaaaaa

FIG. 1B

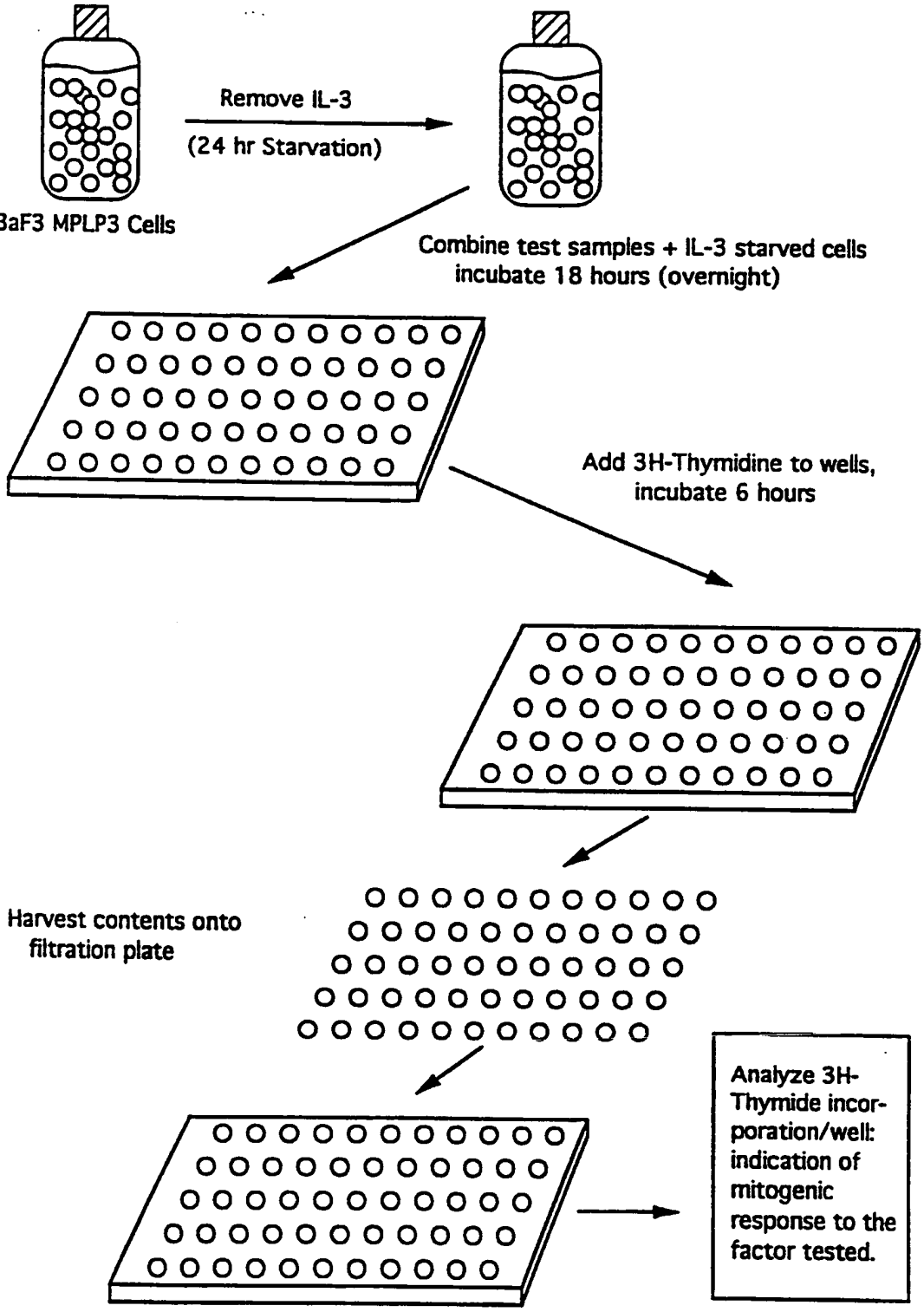

Remove IL-3
(24 hr Starvation)
BaF3 MPLP3 Cells
Combine test samples + IL-3 starved cells
incubate 18 hours (overnight)
Add 3H-Thymidine to wells,
incubate 6 hours
Harvest contents onto
filtration plate
Analyze 3H-
Thymide incor-
poration/well:
indication of
mitogenic
response to the
factor tested.

FIG.2

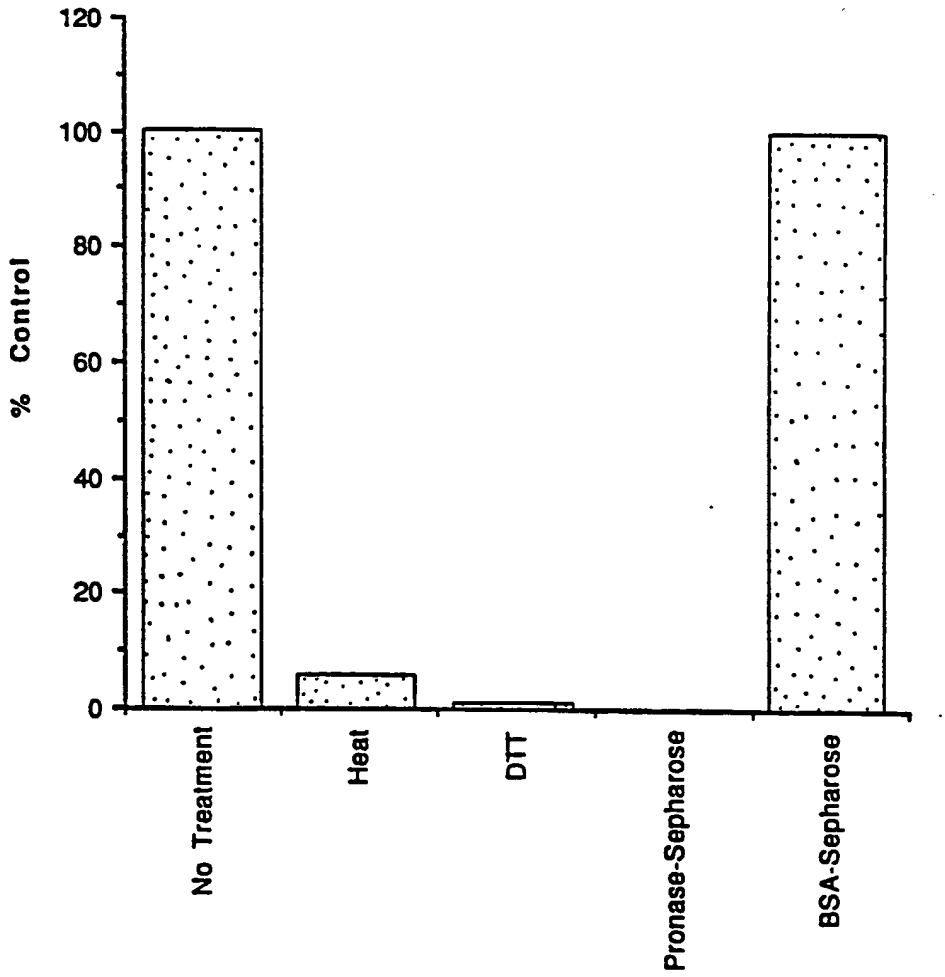

120
100
80
60
40
20
0
% Control
No Treatment
Heat
DTT
Pronase-Sepharose
BSA-Sepharose

FIG.3

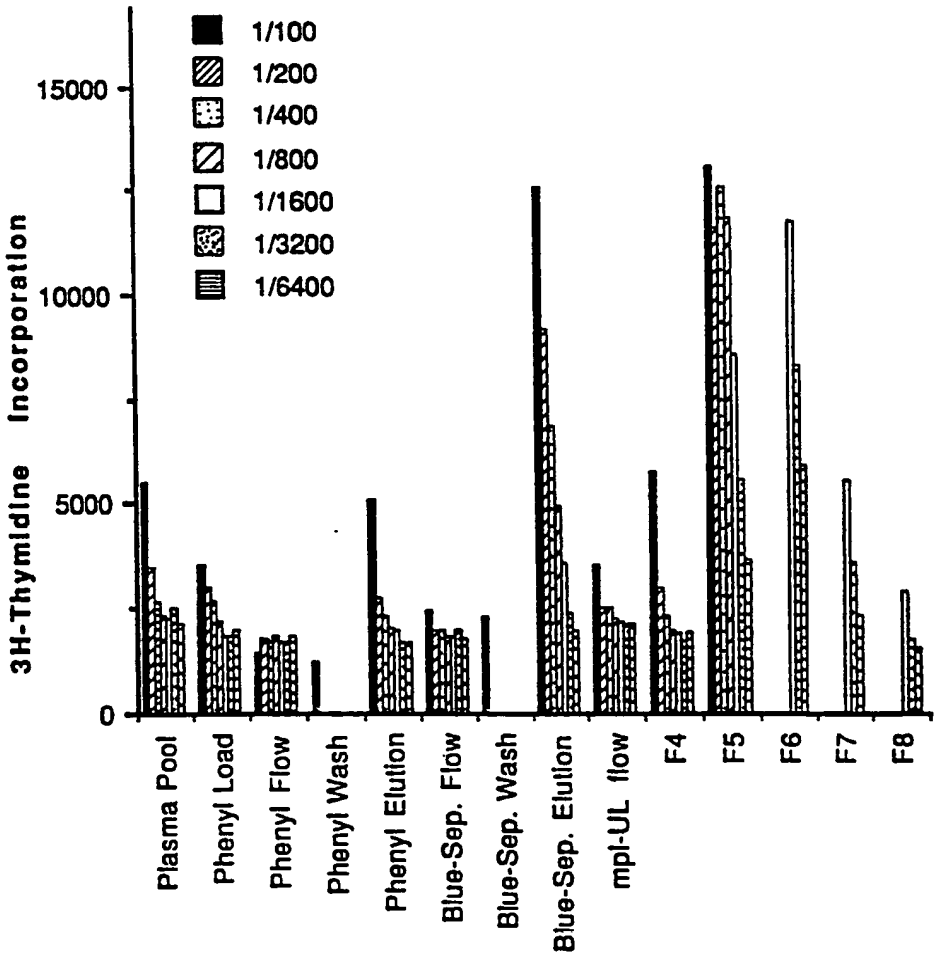

1/100
1/200
1/400
1/800
1/1600
1/3200
1/6400
3H-Thymidine Incorporation
15000
10000
5000
0
Plasma Pool
Phenyl Load
Phenyl Flow
Phenyl Wash
Phenyl Elution
Blue-Sep. Flow
Blue-Sep. Wash
Blue-Sep. Elution
mpl-UL flow
F4
F5
F6
F7
F8

FIG.4

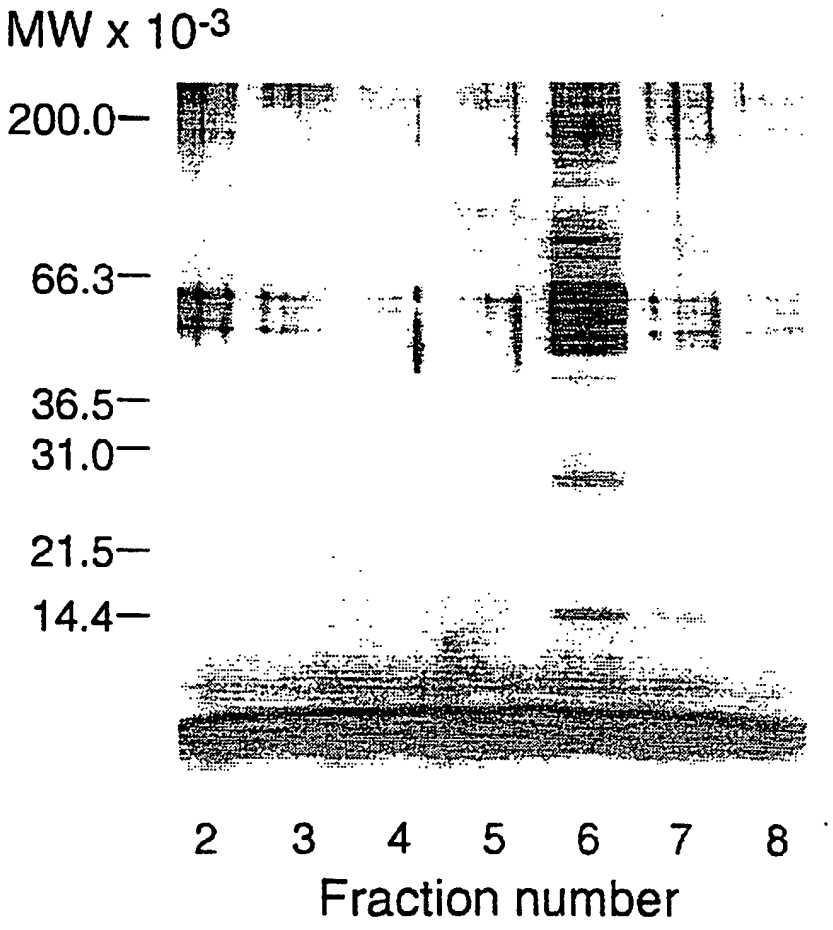


FIG. 5

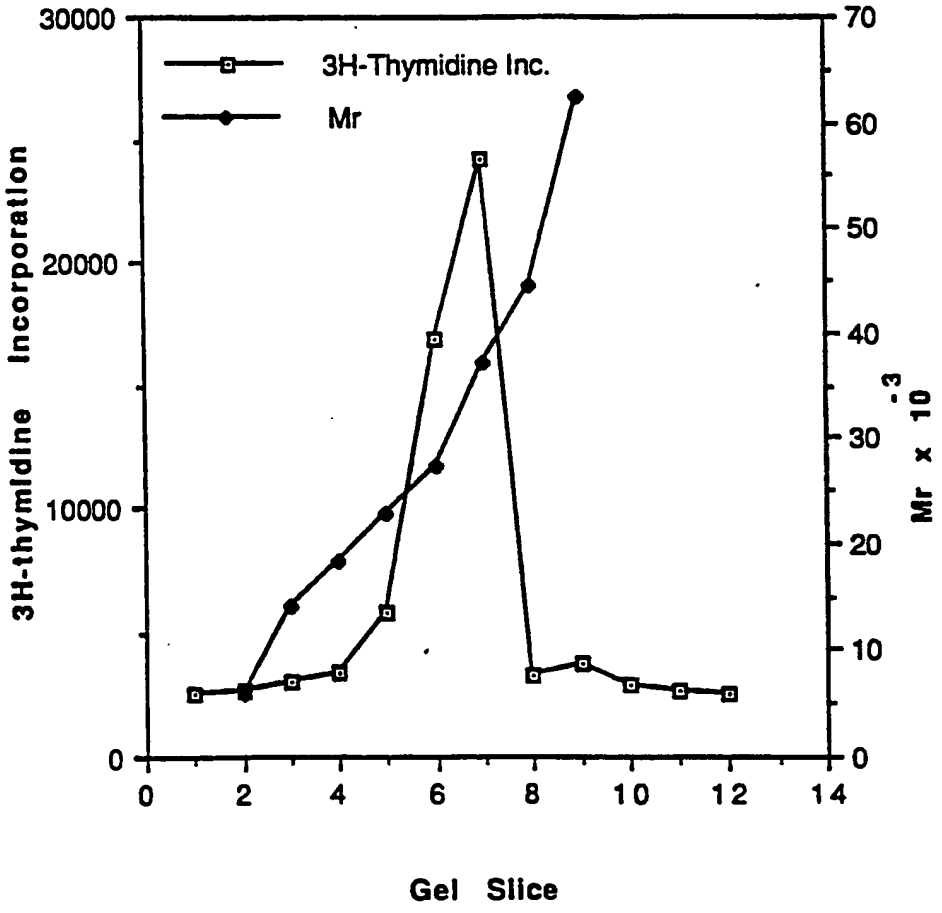

30000
20000
10000
0
3H-Thymidine Inc.
Mr
3H-thymidine Incorporation
70
60
50
40
30
20
10
0
Mr x 10 -3
0
2
4
6
8
10
12
14
Gel Slice

FIG.6

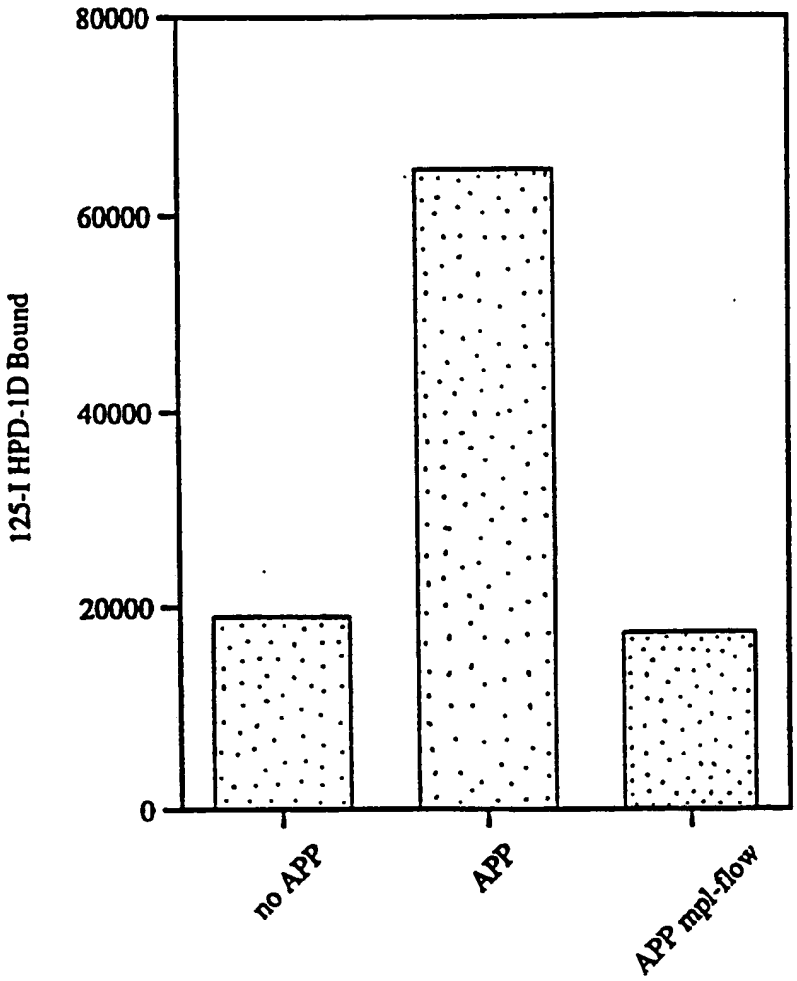

80000
60000
40000
20000
0
125-I HPD-1D Bound
no APP
APP
APP mpl-flow

FIG.7

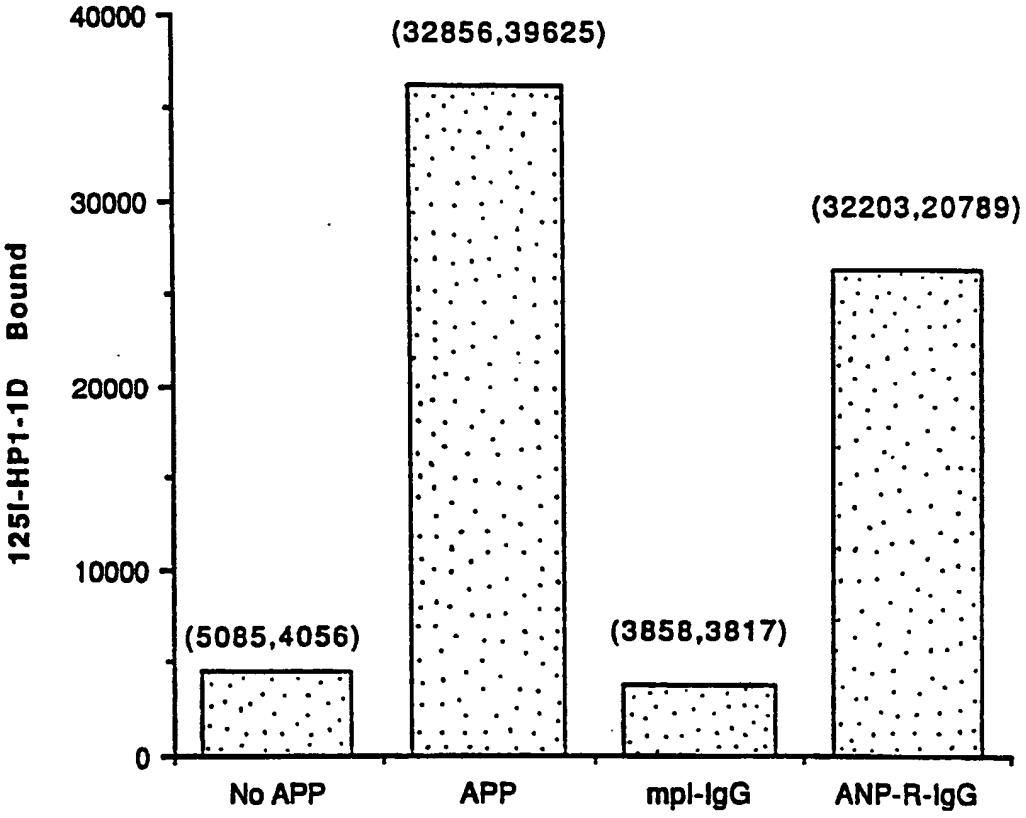

40000
30000
20000
10000
0
125I-HP1-1D Bound
(32856,39625)
(32203,20789)
(5085,4056)
(3858,3817)
No APP
APP
mpl-IgG
ANP-R-IgG

FIG.8

```
                                                                                               -10
                                                                          ↓ L  L  L  V   V  M  L   L  L  T
  1 GAATTCCTGG AATACCAGCT GACAATGATT TCCTCCTCAT CTTTCAACCT CACCTCTCCT CATCTAAGAA TTGCTCCTCG TGGTCATGCT TCTCCTAACT
    CTTAAGGACC TTATGGTCGA CTGTTACTAA AGGAGGAGTA GAAAGTTGGA GTGGAGAGGA GTAGATTCTT AACGAGGAGC ACCAGTACGA AGAGGATTGA

                                                   10                               20
     A  R  L  T   L  S  S   P  A  P   P  A  C  D  L  R  V   L  S  K   L  L  R  D  S  H  V   L  H  S   R  L ↓
101 GCAAGGCTAA CGCTGTCCAG CCCGGCTCCT CCTGCTTGTG ACCTCCGAGT CCTCAGTAAA CTGCTTCGTG ACTCCCATGT CCTTCACAGC AGACTGGTGA
    CGTTCCGATT GCGACAGGTC GGGCCGAGGA GGACGAACAC TGGAGGCTCA GGAGTCATTT GACGAAGCAC TGAGGGTACA GGAAGTGTCG TCTGACCACT

201 GAACTCCCAA CATTATCCCC TTTATCCGCG TAACTGGTAA GACACCCATA CTCCCAGGAA GACACCATCA CTTCCTCTAA CTCCTTGACC CAATGACTAT
    CTTGAGGGTT GTAATAGGGG AAATAGGCGC ATTGACCATT CTGTGGGTAT GAGGGTCCTT CTGTGGTAGT GAAGGAGATT GAGGAACTGG GTTACTGATA

301 TCTTCCCATA TTGTCCCCAC CTACTGATCA CACTCTCTGA CAAGAATTAT TCTTCACAAT ACAGCCCGCA TTTAAAAGCT CTCGTCTAGA
    AGAAGGGTAT AACAGGGGTG GATGACTAGT GTGAGAGACT GTTCTTAATA AGAAGTGTTA TGTCGGGCGT AAATTTTCGA GAGCAGATCT
```

FIG.9

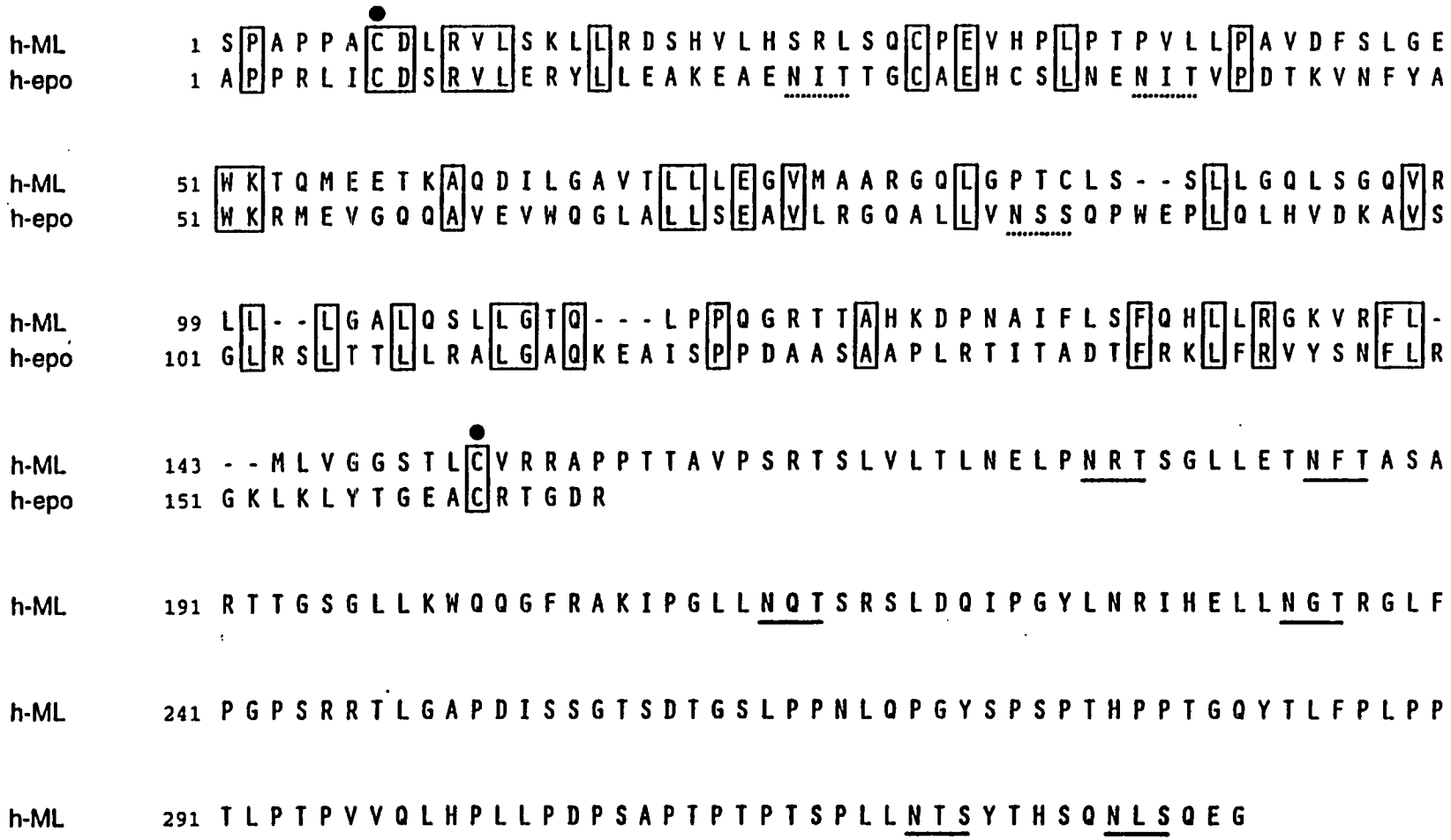

h-ML 1 S P A P P A C D L R V L S K L L R D S H V L H S R L S Q C P E V H P L P T P V L L P A V D F S L G E
h-epo 1 A P P R L I C D S R V L E R Y L L E A K E A E N I T T G C A E H C S L N E N I T V P D T K V N F Y A
h-ML 51 W K T Q M E E T K A Q D I L G A V T L L L E G V M A A R G Q L G P T C L S - - S L L G Q L S G Q V R
h-epo 51 W K R M E V G Q Q A V E V W Q G L A L L S E A V L R G Q A L L V N S S Q P W E P L Q L H V D K A V S
h-ML 99 L L - - L G A L Q S L L G T Q - - - L P P Q G R T T A H K D P N A I F L S F Q H L L R G K V R F L -
h-epo 101 G L R S L T T L L R A L G A Q K E A I S P P D A A S A A P L R T I T A D T F R K L F R V Y S N F L R
h-ML 143 - - M L V G G S T L C V R R A P P T T A V P S R T S L V L T L N E L P N R T S G L L E T N F T A S A
h-epo 151 G K L K L Y T G E A C R T G D R
h-ML 191 R T T G S G L L K W Q Q G F R A K I P G L L N Q T S R S L D Q I P G Y L N R I H E L L N G T R G L F
h-ML 241 P G P S R R T L G A P D I S S G T S D T G S L P P N L Q P G Y S P S P T H P P T G Q Y T L F P L P P
h-ML 291 T L P T P V V Q L H P L L P D P S A P T P T P T S P L L N T S Y T H S Q N L S Q E G

FIG. 10

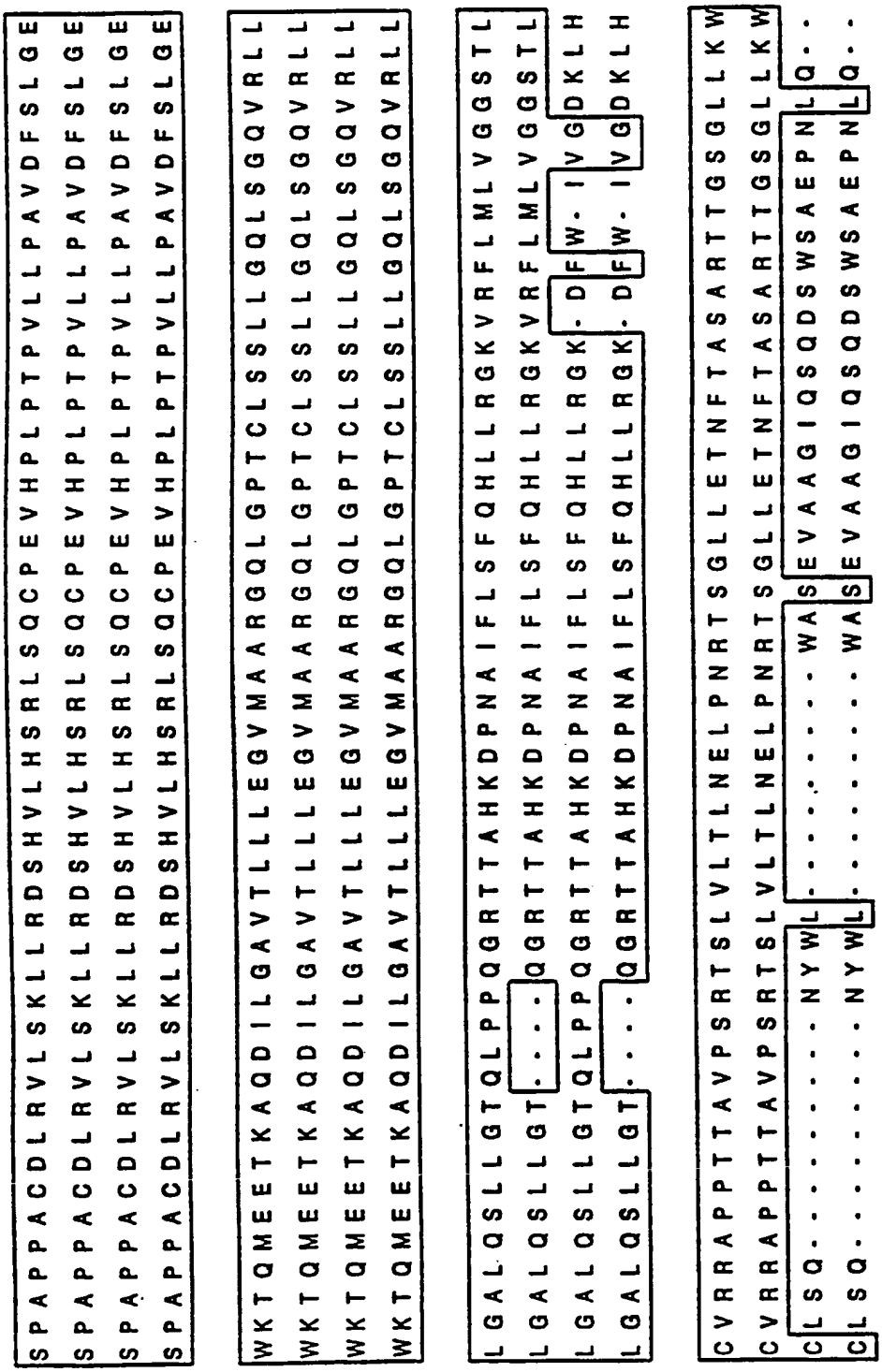

hML 1 SPAPPACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGE
hML2 1 SPAPPACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGE
hML3 1 SPAPPACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGE
hML4 1 SPAPPACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGE
hML 51 WKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLL
hML2 51 WKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLL
hML3 51 WKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLL
hML4 51 WKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLL
hML 101 LGALQSLLGTQLPPQGRTTAHKDPNAIFLSFQHLLRGKVRFLMLVGGSTL
hML2 101 LGALQSLLGT----QGRTTAHKDPNAIFLSFQHLLRGKVRFLMLVGGSTL
hML3 101 LGALQSLLGTQLPPQGRTTAHKDPNAIFLSFQHLLRGK-DFW-IVGDKLH
hML4 101 LGALQSLLGT----QGRTTAHKDPNAIFLSFQHLLRGK-DFW-IVGDKLH
hML 151 CVRRAPPTTAVPSRTSLVLTLNELPNRTSGLLETNFTASARTTGSGLLKW
hML2 147 CVRRAPPTTAVPSRTSLVLTLNELPNRTSGLLETNFTASARTTGSGLLKW
hML3 149 CLSQ----------NYWL----------WASEVAAGIQSQDSWSAEPNLQ--
hML4 145 CLSQ----------NYWL----------WASEVAAGIQSQDSWSAEPNLQ--

FIG. 11A

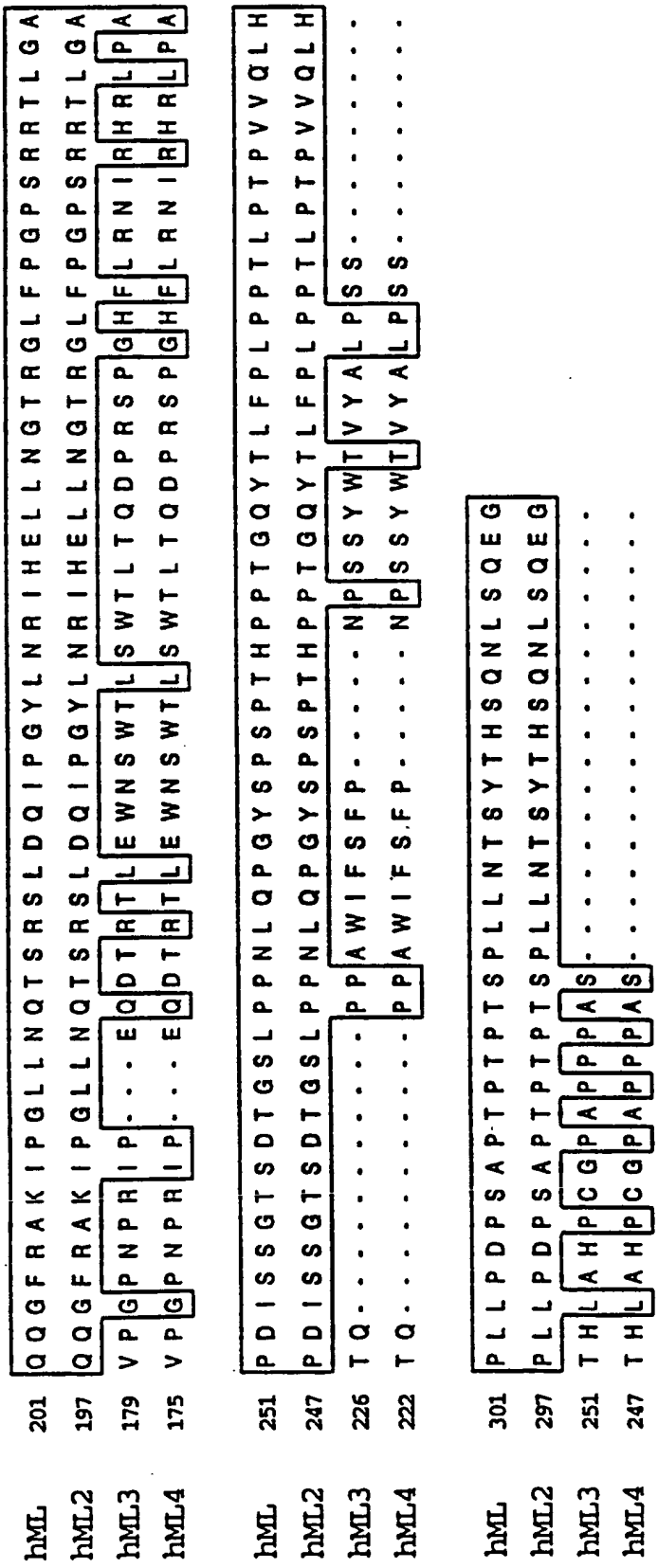

hML 201 QQGFRAKIPGLLNQTSRSLDQIPGYLNRIHELLNGTRGLFPGPSRRTLGA
hML2 197 QQGFRAKIPGLLNQTSRSLDQIPGYLNRIHELLNGTRGLFPGPSRRTLGA
hML3 179 VPGPNPRIP---EQDTRTLEWNSWTLSWTLTQDPRSPGHFLRNIRHRLPA
hML4 175 VPGPNPRIP---EQDTRTLEWNSWTLSWTLTQDPRSPGHFLRNIRHRLPA
hML 251 PDISSGTSDTGSLPPNLQPGYSPSPTHPPTGQYTLFPLPPTLPTPVVQLH
hML2 247 PDISSGTSDTGSLPPNLQPGYSPSPTHPPTGQYTLFPLPPTLPTPVVQLH
hML3 226 TQ----------PPAWIFSFP-----NPSSYWTVYALPSS----------
hML4 222 TQ----------PPAWIFSFP-----NPSSYWTVYALPSS----------
hML 301 PLLPDPSAPTPTPTSPLLNTSYTHSQNLSQEG
hML2 297 PLLPDPSAPTPTPTSPLLNTSYTHSQNLSQEG
hML3 251 THLAHPCGPAPPPAS----------
hML4 247 THLAHPCGPAPPPAS----------

FIG. 11B

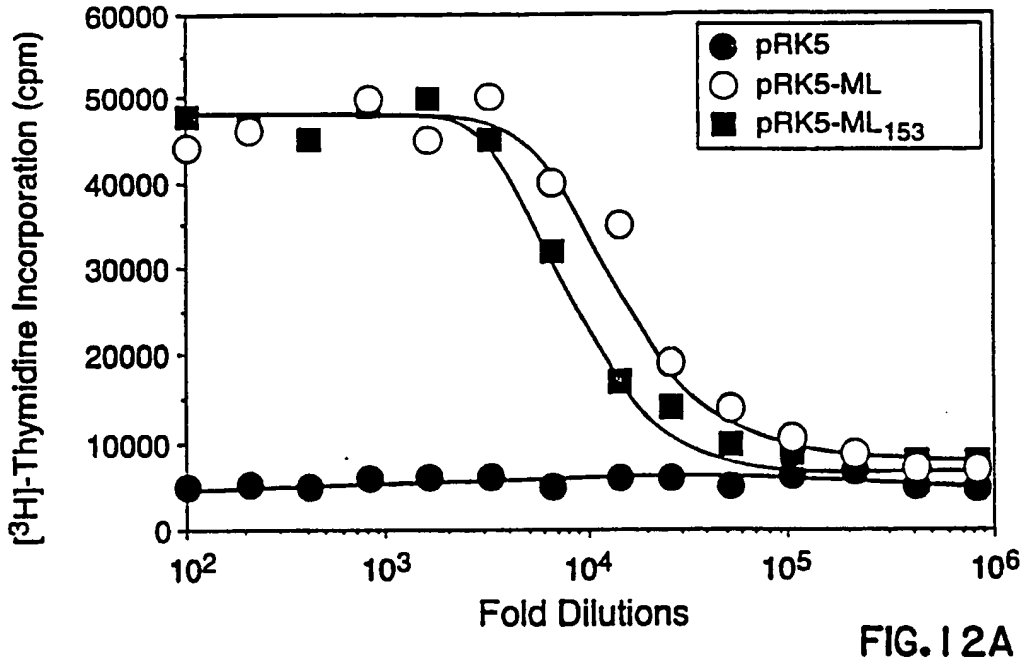


FIG. 12A

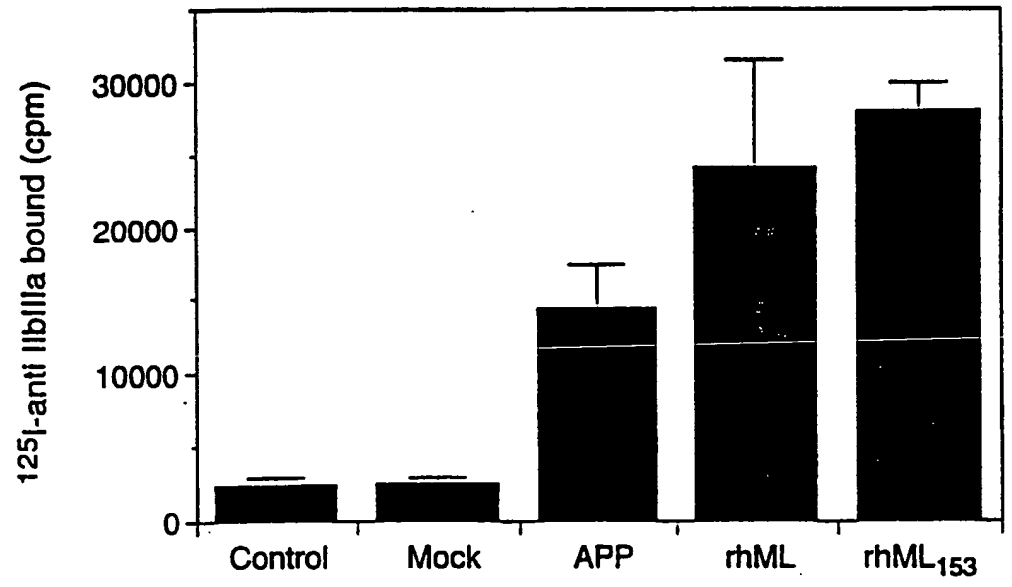


FIG. 12B

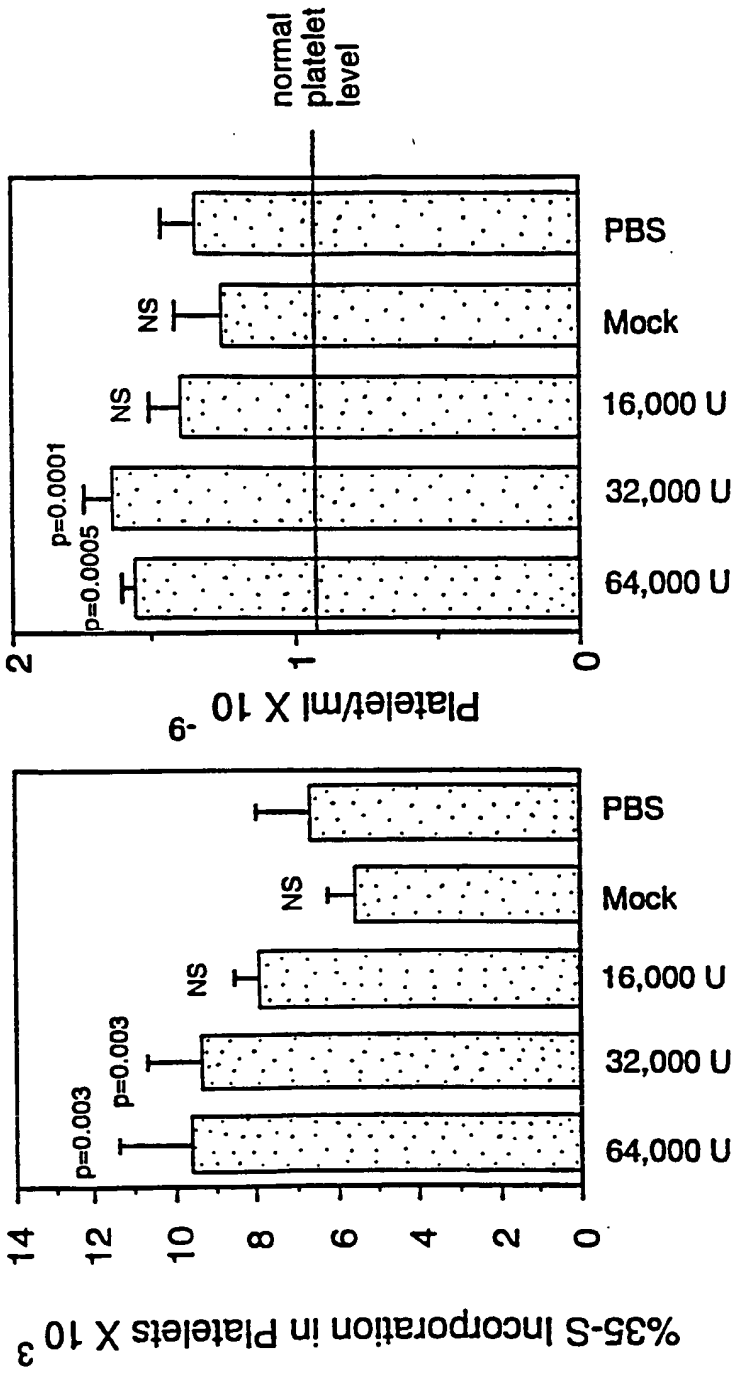

normal platelet level
PBS
Mock
16,000 U
32,000 U
64,000 U
NS
NS
p=0.0001
p=0.0005
2
1
0
Platelet/ml X 10 -9
PBS
Mock
16,000 U
32,000 U
64,000 U
NS
NS
p=0.003
p=0.003
14
12
10
8
6
4
2
0
%35-S Incorporation in Platelets X 10 3

FIG. 12C

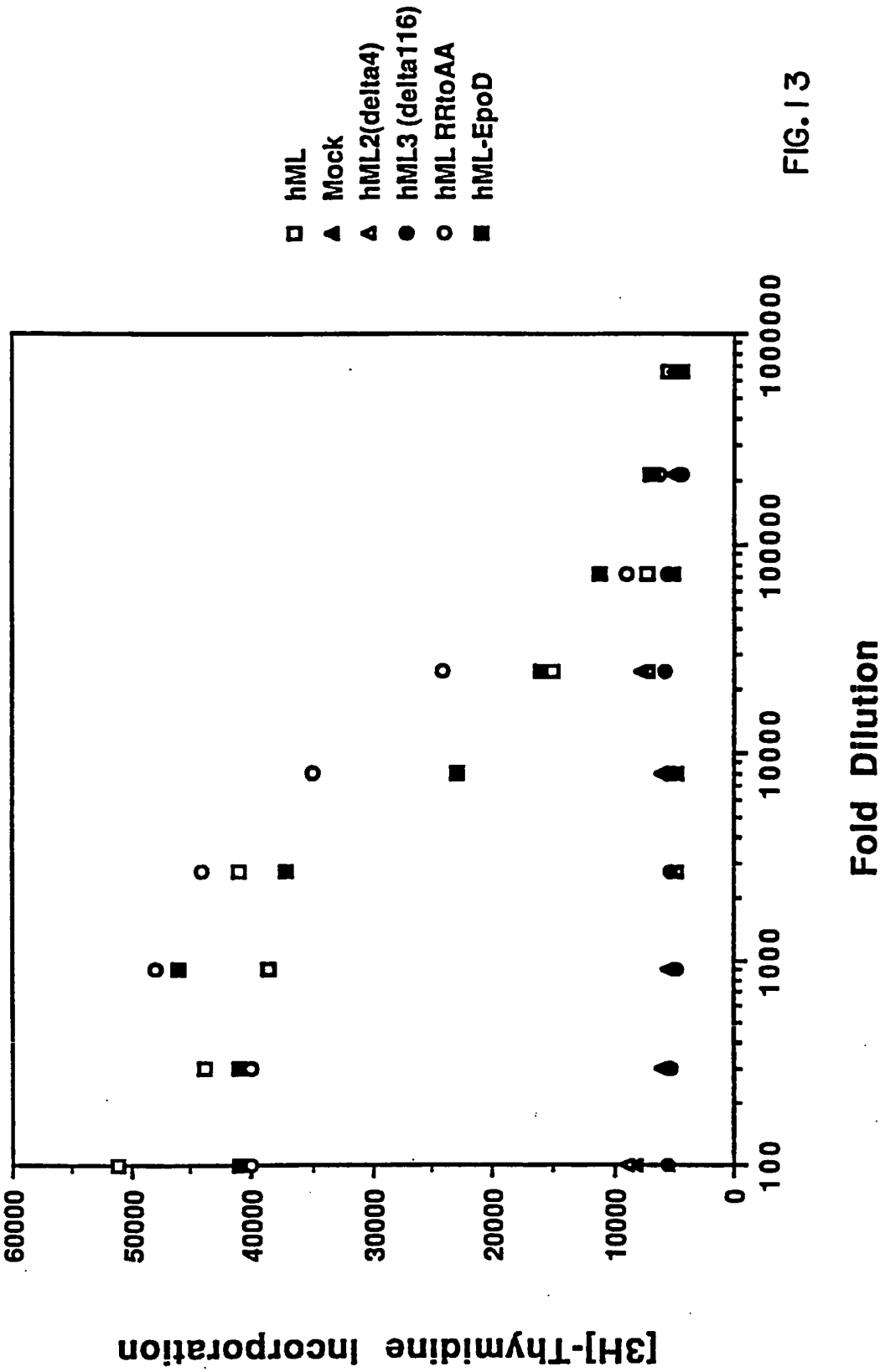

hML
Mock
hML2(delta4)
hML3 (delta116)
hML RRtoAA
hML-EpoD
[3H]-Thymidine Incorporation
60000
50000
40000
30000
20000
10000
0
100
1000
10000
100000
1000000
Fold Dilution

FIG. 13

```
   1 cccagcctcctttctcttgttccctggtcatgcctgcctccctgtctcctgtctctccctcccacacacacccactatcctcccagctatccctacaccc
 101 tccttcctaatcttgggagacatctcgtctggctggacgggaaaattccaggatctaggccacacttctcagcagacatgcccatccttggggaggagga
 201 acaggagagagcctgaggaagttctgggggacaggggatgatgggatcaaggtcaggccaggaagcccctgaggacagagactgtggggagactgggac
 301 tgggaagaaagcaaaggagctagagccagggccaaaggaaaaggggggccagcagggaggtatttgcgggggaggtccagcagctgtctttcctaagaca
 401 gggacacatgggcctggttattcctcttgtcacatgtggaacggtaggagatggaagacggagacagaacaagcaaaggagggccctgggcacagaggtc
 501 tgtgtgtgtagccatccaagccactggaccccagcagacgagcacctaagctcaggcttaacccagtgcacgtgtgcgcacatacatgtgccccgcacct
 601 gacagtccactcaacccgtccaaacccttttccccataacaccaacccataacaggagatttctctcatgtgggcaatatccgtgttcccacttcgaaagg
 701 gggaatgacaagataggactccctaggggattacagaaagaaaagcaggaaagcaagcatcctgttggatttcagcagcaggtatgatgtccagggaaaa
 801 gaaatttggatagccagggagtgaaaaccccaccaatcttaaacaagacctctgtgcttcttccccagcaacacaaatgtcctgccagattcctcctgga
 901 aaaaaacttctgctcctgtccccctccaggtcccaggttgcccatgtccaggaaaagatggatccccctatccaaatcttctccgtggtgtgtgtgggtg
1001 gaggagtggaccctggtccaggcaggggctccagggaagagaaggcgtcacttccgggggccttcaccagtgtctggtggctcccttctctgattgggca
                                                                    < Start of cDNA sequence; Exon 1 >
1101 gaagtggcccaggcaggcgtatgacctgctgctgtggaggggctgtgccccaccgccacatgtctTCCTACCCATCTGCTCCCCAGAGGGCTGCCTGCTG
1201 TGCACTTGGGTCCTGGAGCCCTTCTCCACCCGgtgagtggccagcagggtgtggggttatgtgagggtagaaaggacagcaaagagaaatgggctcccag
1301 ctgggggaggggcaggcaaactggaacctacaggcactgacctttgtcgagaagagtgtagccttcccagaatgggaggagcagggcagagcaggggtag
1401 ggggtggggtgctggtttctgagggactgatcacttacttggtggaatacagcacagccctggctggccctaaggaaaggggacatgagcccagggagaa
1501 aataagagagggagctgcacttagggcttagcaaacacagtagtaagatggacacagccccaatccccattcttagctggtcattcctcgttagcttaag
1601 gttctgaatctggtgctggggaagctgggccaggcaagccagggcgcaaggagagggtaatgggaggagggcccactcatgttgacagacctacaggaaa
```

FIG. 14A

```
1701 tcccaatattgaatcaggtgcaagcctctttgcacaacttgtgaaaggaggaggaagccatgtgggggtcctgtgaaggaaccggaaggggttctgcca

1801 agggggcagggaggcaggtgtgagctatgagacagatatgttagtgggcgcctaagacaaggtaagcccctaaggtgggcatcacccagcaggtgcccgt

1901 tcctgggcagctggtctcaggaaggaagtcccagaactgttagcccatctcttggcctcagataatggagtatttcaggacttggagtccaaagaaaagc

2001 tccagtggctttatgtgtgggggtagatagggaaagaatagaggttaatttctcccataccgccttttaatcctgacctctagtggtcccagttacagct

2101 ttgtgcagttcccctccccagccccactccccaccgcagaagttacccctcaacatattgcgcccgtttgccagttcctcacccaggccctgcatcccat

2201 tttccactctcttctccaggctgaagccacaatactttccttctctatccccatcccagattttctctgacctaacaaccaaggttgctcagaatttaag

2301 gctaattaagatatgtgtgtatacatatcatgtcctgctgctctcagcaggggtaggtggcaccaaatccgtgtccgattcactgaggagtcctgacaaa

2401 aaggagacaccatatgctttcttgctttctttctttctttctttctttttttttttttgagacggagtttcactcttattgcccaggctggagtgcaatg

2501 gtgcgatctcggctcaccacaaacctccgcctcccaggtacaagcgattctcctgtctcagcctcccaagtagcttggattacaggcatgagccaccaca

2601 ccctgctagtttttttgtatttcgtagagccggggtttcaccatgttagtgaggctggtggcgaactcctgacctcaggtgatccacccgccttggactc

2701 ccaaagtgctgggattacaggcatgagccactgcacccggcacaccatatgctttcatcacaagaaaatgtgagagaattcagggctttggcagttccag

2801 gctggtcagcatctcaagccctccccagcatctgttcaccctgccaggcagtctcttcctagaaacttggttaaatgttcactcttcttgctactttcag

     < Exon 2 >
2901 GATAGATTCCTCACCCTTGGCCCGCCTTTGCCCCACCCTACTCTGCCCAGAAGTGCAAGAGCCTAAGCCGCCTCCATGGCCCCAGGAAGGATTCAGGGGA

   1                                                MetGluLeuThrG
3001 GAGGCCCCAAACAGGGAGCCACGCCAGCCAGACACCCCGGCCAGAATGGAGCTGACTGgtgagaacacacctgaggggctagggccatatggaaacatga

3101 cagaaggggagagagaaaggagacacgctgcaggggcaggaagctgggggaacccattctcccaaaaataaggggtctgaggggtggattccctgggtt
                                                                                    < Exon 3 >
   6                                                                                luLeuLeuLeu
3201 tcaggtctgggtcctgaatgggaattcctggaataccagctgacaatgatttcctcctcatctttcaacctcacctctcctcatctaagAATTGCTCCTC
```

FIG. 14B

```
   9 ValValMetLeuLeuLeuThrAlaArgLeuThrLeuSerSerProAlaProProAlaCysAspLeuArgValLeuSerLysLeuLeuArgAspSerHisVal
3301 GTGGTCATGCTTCTCCTAACTGCAAGGCTAACGCTGTCCAGCCCGGCTCCTCCTGCTTGTGACCTCCGAGTCCTCAGTAAACTGCTTCGTGACTCCCATG
                    End of signal peptide^

  43   LeuHisSerArgLeu
3401 TCCTTCACAGCAGACTGgtgagaactcccaacattatccccttttatccgcgtaactggtaagacacccatactcccaggaagacaccatcacttcctcta

3501 actccttgacccaatgactattcttcccatattgtccccacctactgatcacactctctgacaagaattattcttcacaatacagcccgcatttaaaagc

3601 tctcgtctagagatagtactcatggaggactagcctgcttattaggctaccatagctctctctatttcagctcccttctccccccaccaatctttttcaa
        < Exon 4 >
  48    SerGlnCysProGluValHisProLeuProThrProValLeuLeuProAlaValAspPheSerLeuGlyGluTrpLysThrGlnMet
3701 cagAGCCAGTGCCCAGAGGTTCACCCTTTGCCTACACCTGTCCTGCTGCCTGCTGTGGACTTTAGCTTGGGAGAATGGAAAACCCAGATGgtaagaaagc

3801 catccctaaccttggcttccctaagtcctgtcttcagtttcccactgcttcccatggattctccaacattcttgagctttttaaaaatatctcaccttca

3901 gcttggccaccctaacccaatctacattcacctatgatgatagcctgtggataagatgatggcttgcaggtccaatatgtgaatagatttgaagctgaac

4001 accatgaaaagctggagagaaatcgctcatggccatgcctttgacctattccygttcagtcttcttaaattggcatgaagaagcaagactcatatgtcat

4101 ccacagatgacacaaagctgggaagtaccactaaaataacaaaagactgaatcaagattcaaatcactgaaagactaggtcaaaaacaaggtgaaacaac

4201 agagatataaacttctacatgtgggccgggggctcacgcctgtaatcccagcactttgggaggccgaggcaggcagatcacctgagggcaggagtttgag

4301 agcagcctggccaacatggcgaaaccccgtctctactaagaatacaaaattagccgggcatggtagtgcatgcctgtaatcccagctacttggaaggctg

4401 aagcaggagaatcccttgaacccaggaggtggaggttgtagtgagctgagatcatgccaatgcactccagcctgggtgacaagagcaaaactccgtctca

4501 aaaagaaaaaaaaattctacatgtgtaaattaatgagtaaagtcctattccagctttcaggccacaatgccctgcttccatcatttaagcctctggccct

4601 agcacttcctacgaaaaggatctgagagaattaaattgcccccaaacttaccatgtaacattactgaagctgctattcttaaagctagtaattcttgtct

4701 gtttgatgtttagcatccccattgtggaaatgctcgtacagaactctattccgagtggactacacttaaatatactggcctgaacaccggacatcccccct

4801 gaagacatatgctaatttattaagagggaccatattaaactaacatgtgtctagaaagcagcagcctgaacagaaagagactagaagcatgttttatggg
```

FIG. 14C

```
4901 caatagtttaaaaaactaaaatctatcctcaagaaccctagcgtcccttcttccttcaggactgagtcagggaagaagggcagttcctatgggtcccttc

5001 tagtcctttcttttcatccttatgatcattatggtagagtctcatacctacatttagtttatttattattattatttgagacggagtctcactctatccc

5101 ccaggctggagtgcagtggcatgatctcaactcactgcaacctcagcctcccggattcaagcgattctcctgcctcagtctcccaagtagctgggattac

5201 aggtgcccaccaccatgcccagctaatttttgtatttttggtagagatggggtttcaccatgttggccaggctgatcttgaactcctgacctcaggtgat

5301 ccacctgcctcagcctcccaaagtgctgggattacaggcgtgagccactgcacccagccttcattcagtttaaaaatcaaatgatcctaaggttttgcag

5401 cagaaagagtaaatttgcagcactagaaccaagaggtaaaagctgtaacagggcagatttcagcaacgtaagaaaaaaggagctcttctcactgaaacca

5501 agtgtaagaccaggctggactagaggacacgggagtttttgaagcagaggctgatgaccagctgtcgggagactgtgaaggaattcctgccctgggtggg

5601 accttggtcctgtccagttctcagcctgtatgattcactctgctggctactcctaaggctccccacccgcttttagtgtgccctttgaggcagtgcgctt
                          < Exon 5 >
  77                      GluGluThrLysAlaGlnAspIleLeuGlyAlaValThrLeuLeuLeuGluGlyValMetAlaAlaArgGlyGlnLeu
5701 ctctcttccatctctttctcagGAGGAGACCAAGGCACAGGACATTCTGGGAGCAGTGACCCTTCTGCTGGAGGGAGTGATGGCAGCACGGGGACAACTG

 103 GlyProThrCysLeuSerSerLeuLeuGlyGlnLeuSerGlyGlnValArgLeuLeuLeuGlyAlaLeuGlnSerLeuLeuGlyThrGln
5801 GGACCCACTTGCCTCTCATCCCTCCTGGGGCAGCTTTCTGGACAGGTCCGTCTCCTCCTTGGGGCCCTGCAGAGCCTCCTTGGAACCCAGgtaagtcccc

5901 agtcaagggatctgtagaaactgttcttttctgactcagtcccactagaagacctgagggaagaagggctcttccagggagctcaagggcagaagagctg

6001 atctactaagagtgctccctgccagccacaatgcctgggtactggcatcctgtctttcctacttagacaagggaggcctgagatctggccctggtgtttg
                          < Exon 6 >
 133                      LeuProProGlnGlyArgThrThrAlaHisLysAspProAsnAlaIlePheLeuSerPheGlnHisLeuLeuArg
6101 GCCTCAGGACCATCCTCTGCCCTCAGCTTCCTCCACAGGGCAGGACCACAGCTCACAAGGATCCCAATGCCATCTTCCTGAGTTTCCAACACCTGCTCCG
                          ============^Alternative splice site                    ^C for cDNA clone

 158  GlyLysValArgPheLeuMetLeuValGlyGlySerThrLeuCysValArgArgAlaProProThrThrAlaValProSerArgThrSerLeuValLeu
6201 AGGAAAGGTGCGTTTCCTGATGCTTGTAGGAGGGTCCACCCTCTGCGTCAGGCGGGCCCCACCCACCACAGCTGTCCCCAGCAGAACCTCTCTAGTCCTC
                                                  ^End of EPO dormain
 191 ThrLeuAsnGluLeuProAsnArgThrSerGlyLeuLeuGluThrAsnPheThrAlaSerAlaArgThrThrGlySerGlyLeuLeuLysTrpGlnGlnGly
6301 ACACTGAACGAGCTCCCAAACAGGACTTCTGGATTGTTGGAGACAAACTTCACTGCCTCAGCCAGAACTACTGGCTCTGGGCTTCTGAAGTGGCAGCAGG
```

FIG. 14D

```
 225   PheArgAlaLysIleProGlyLeuLeuAsnGlnThrSerArgSerLeuAspGlnIleProGlyTyrLeuAsnArgIleHisGluLeuLeuAsnGlyThr
6401 GATTCAGAGCCAAGATTCCTGGTCTGCTGAACCAAACCTCCAGGTCCCTGGACCAAATCCCCGGATACCTGAACAGGATACACGAACTCTTGAATGGAAC

 258  ArgGlyLeuPheProGlyProSerArgArgThrLeuGlyAlaProAspIleSerSerGlyThrSerAspThrGlySerLeuProProAsnLeuGlnPro
6501 TCGTGGACTCTTTCCTGGACCCTCACGCAGGACCCTAGGAGCCCCGGACATTTCCTCAGGAACATCAGACACAGGCTCCCTGCCACCCAACCTCCAGCCT

 291 GlyTyrSerProSerProThrHisProProThrGlyGlnTyrThrLeuPheProLeuProProThrLeuProThrProValValGlnLeuHisProLeuLeu
6601 GGATATTCTCCTTCCCCAACCCATCCTCCTACTGGACAGTATACGCTCTTCCCTCTTCCACCCACCTTGCCCACCCCTGTGGTCCAGCTCCACCCCCTGC

 325   ProAspProSerAlaProThrProThrProThrSerProLeuLeuAsnThrSerTyrThrHisSerGlnAsnLeuSerGlnGluGly:STOP
6701 TTCCTGACCCTTCTGCTCCAACGCCCACCCCTACCAGCCCTCTTCTAAACACATCCTACACCCACTCCCAGAATCTGTCTCAGGAAGGGTAAGGTTCTCA

6801 GACACTGCCGACATCAGCATTGTCTCATGTACAGCTCCCTTCCCTGCAGGGCGCCCCTGGGAGACAACTGGACAAGATTTCCTACTTTCTCCTGAAACCC

6901 AAAGCCCTGGTAAAAGGGATACACAGGACTGAAAAGGGAATCATTTTTCACTGTACATTATAAACCTTCAGAAGCTATTTTTTTAAGCTATCAGCAATAC

7001 TCATCAGAGCAGCTAGCTCTTTGGTCTATTTTCTGCAGAAATTTGCAACTCACTGATTCTCTACATGCTCTTTTTCTGTGATAACTCTGCAAAGGCCTGG

7101 GCTGGCCTGGCAGTTGAACAGAGGGAGAGACTAACCTTGAGTCAGAAAACAGAGAAAGGGTAATTTCCTTTGCTTCAAATTCAAGGCCTTCCAACGCCCC

7201 CATCCCCTTTACTATCATTCTCAGTGGGACTCTGATCCCATATTCTTAACAGATCTTTACTCTTGAGAAATGAATAAGCTTTCTCTCAGaaatgctgtcc
                                                                                                ^PolyA site

7301 ctatacactagacaaaactgagcctgtataaggaataaatgggagcgccgaaaagctccctaaaaagcaagggaaagatgttcttcgagggtggcaatag

7401 atccccctcaccctgccaccccaaacaaaaaagctaacaggaagccttggagagcctcacaccccaggtaaggctgtgtagacagttcagtaaagacagg

7501 acctggatgtgacagctgagcaaacagctagagctttggcagctcagcaggaggctttgccaggcatggacgcctgcctccctcctgtggaggtcaggag

7601 gaagtgcaggaagtggcatgagtcaggctccttgagctcacacagcaggagaacaagtacaagtcaagtacaagttgaaggctcatttcccagttcccgc

7701 aaatgcatctaaaaagcagctctgtgtgaccaccataaactctgctaggggatctctaaaaaggagtcaggcttatggggctttgcaaataagtgctgcc

7801 ttggtgctcaggaaaaggtttgtgttgcacaaaacacaaattccactgc
```

FIG. 14E

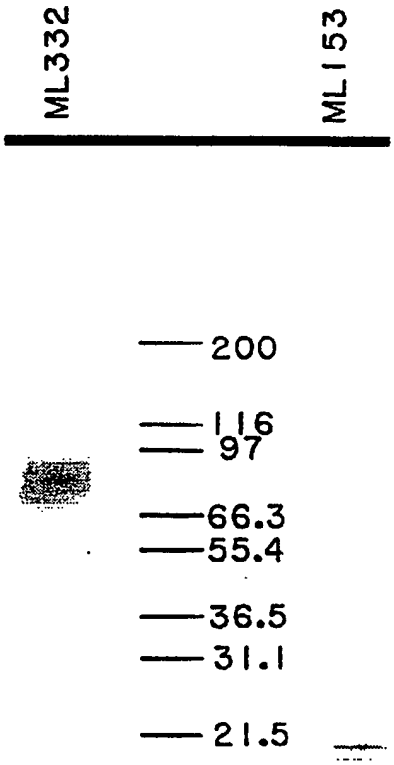


FIG. 15

```
  1 GAGTCCTTGG CCCACCTCTC TCCCACCCGA CTCTGCCGAA AGAAGCACAG AAGCTCAAGC CGCCTCCATG GCCCCAGGAA AGATTCAGGG GAGAGGCCCC

                                                 -20                               -10
                                            Met GluLeuThrA spLeuLeuLe uAlaAlaMet LeuLeuAlaV alAlaArgLe uThrLeuSer
101 ATACAGGGAG CCACTTCAGT TAGACACCCT GGCCAGAATG GAGCTGACTG ATTTGCTCCT GGCGGCCATG CTTCTTGCAG TGGCAAGACT AACTCTGTCC

                 ●          10                                20                                ● 30
    SerProValA laProAlaCy sAspProArg LeuLeuAsnL ysLeuLeuAr gAspSerHis LeuLeuHisS erArgLeuSe rGlnCysPro AspValAspPro
201 AGCCCCGTAG CTCCTGCCTG TGACCCCAGA CTCCTAAATA AACTGCTGCG TGACTCCCAC CTCCTTCACA GCCGACTGAG TCAGTGTCCC GACGTCGACC

                     40                                50                                60
     LeuSerIl ePro ValLeu LeuProAlaV alAspPheSe rLeuGlyGlu TrpLysThrG lnThrGluGl nSerLysAla GlnAspIleL euGlyAlaVal
301 CTTTGTCTAT CCCTGTTCTG CTGCCTGCTG TGGACTTTAG CCTGGGAGAA TGGAAAACCC AGACGGAACA GAGCAAGGCA CAGGACATTC TAGGGGCAGT

         70                          80                      ●           90                                100
     SerLeuLeu LeuGluGlyV alMetAlaAl aArgGlyGln LeuGluProS erCysLeuSe rSerLeuLeu GlyGlnLeuS erGlyGlnVa lArgLeuLeu
401 GTCCCTTCTA CTGGAGGGAG TGATGGCAGC ACGAGGACAG TTGGAACCCT CCTGCCTCTC ATCCCTCCTG GGACAGCTTT CTGGGCAGGT TCGCCTCCTC

                                110                               120                               130
    LeuGlyAlaL euGlnGlyLe uLeuGlyThr GlnGlyArgT hrThrAlaHi sLysAspPro AsnAlaLeuP heLeuSerLe uGlnGlnLeu LeuArgGlyLys
501 TTGGGGGCCC TGCAGGGCCT CCTAGGAACC CAGGGCAGGA CCACAGCTCA CAAGGACCCC AATGCCCTCT TCTTGAGCTT GCAACAACTG CTTCGGGGAA

                     140                    ●         150                               160
     ValArgPh eLeuLeuLeu ValGluGlyP roThrLeuCy sValArgArg ThrLeuProT hrThrAlaVa lProSerSer ThrSerGlnL euLeuThrLeu
601 AGGTGCGCTT CCTGCTTCTG GTAGAAGGTC CCACCCTCTG TGTCAGACGG ACCCTGCCAA CCACAGCTGT CCCAAGCAGT ACTTCTCAAC TCCTCACACT
```

FIG. 16A

```
                  170                                180                                190                                200
     AsnLysPhe ProAsnArgT hrSerGlyLe uLeuGluThr AsnPheSerV alThrAlaAr gThrAlaGly ProGlyLeuL euSerArgLe uGlnGlyPhe
 701 AAACAAGTTC CCAAACAGGA CTTCTGGATT GTTGGAGACG AACTTCAGTG TCACAGCCAG AACTGCTGGC CCTGGACTTC TGAGCAGGCT TCAGGGATTC

                                  210                                220                                230
     ArgValLysI leThrProGl yGlnLeuAsn GlnThrSerA rgSerProVa lGlnIleSer GlyTyrLeuA snArgThrHi sGlyProVal AsnGlyThrHis
 801 AGAGTCAAGA TTACTCCTGG TCAGCTAAAT CAAACCTCCA GGTCCCCAGT CCAAATCTCT GGATACCTGA ACAGGACACA CGGACCTGTG AATGGAACTC

                       240                                250                                260
      GlyLeuPh eAlaGlyThr SerLeuGlnT hrLeuGluAl aSerAspIle SerProGlyA laPheAsnLy sGlySerLeu AlaPheAsnL euGlnGlyGly
 901 ATGGGCTCTT TGCTGGAACC TCACTTCAGA CCCTGGAAGC CTCAGACATC TCGCCCGGAG CTTTCAACAA AGGCTCCCTG GCATTCAACC TCCAGGGTGG

                  270                                280                                290                                300
      LeuProPro SerProSerL euAlaProAs pGlyHisThr ProPheProP roSerProAl aLeuProThr ThrHisGlyS erProProGl nLeuHisPro
1001 ACTTCCTCCT TCTCCAAGCC TTGCTCCTGA TGGACACACA CCCTTCCCTC CTTCACCTGC CTTGCCCACC ACCCATGGAT CTCCACCCCA GCTCCACCCC

                                  310                                320                                330
     LeuPheProA spProSerTh rThrMetPro AsnSerThrA laProHisPr oValThrMet TyrProHisP roArgAsnLe uSerGlnGlu Thr
1101 CTGTTTCCTG ACCCTTCCAC CACCATGCCT AACTCTACCG CCCCTCATCC AGTCACAATG TACCCTCATC CCAGGAATTT GTCTCAGGAA ACATAGCGCG

1201 GGCACTGGCC CAGTGAGCGT CTGCAGCTTC TCTCGGGGAC AAGCTTCCCC AGGAAGGCTG AGAGGCAGCT GCATCTGCTC CAGATGTTCT GCTTTCACCT

1301 AAAAGGCCCT GGGGAAGGGA TACACAGCAC TGGAGATTGT AAAATTTTAG GAGCTATTTT TTTTTAACCT ATCAGCAATA TTCATCAGAG CAGCTAGCGA

1401 TCTTTGGTCT ATTTTCGGTA TAAATTTGAA AATCACTAAT TCT
```

FIG. 16B


1 gagtccttggcccacctctctcccacccgactctgccgaaagaagcacagaagctcaagccgcctccatggccccaggaaagattcaggggagaggcccc
-20 -10
MetGluLeuThrAspLeuLeuLeuAlaAlaMetLeuLeuAlaValAlaArgLeuThrLeuSer
101 atacagggagccacttcagttagacaccctggccagaATGGAGCTGACTGATTTGCTCCTGGCGGCCATGCTTCTTGCAGTGGCAAGACTAACTCTGTCC
10 20 30
SerProValAlaProAlaCysAspProArgLeuLeuAsnLysLeuLeuArgAspSerHisLeuLeuHisSerArgLeuSerGlnCysProAspValAspPro
201 AGCCCCGTAGCTCCTGCCTGTGACCCCAGACTCCTAAATAAACTGCTGCGTGACTCCCACCTCCTTCACAGCCGACTGAGTCAGTGTCCCGACGTCGACC
40 50 60
LeuSerIleProValLeuLeuProAlaValAspPheSerLeuGlyGluTrpLysThrGlnThrGluGlnSerLysAlaGlnAspIleLeuGlyAlaVal
301 CTTTGTCTATCCCTGTTCTGCTGCCTGCTGTGGACTTTAGCCTGGGAGAATGGAAAACCCAGACGGAACAGAGCAAGGCACAGGACATTCTAGGGGCAGT
70 80 90 100
SerLeuLeuLeuGluGlyValMetAlaAlaArgGlyGlnLeuGluProSerCysLeuSerSerLeuLeuGlyGlnLeuSerGlyGlnValArgLeuLeu
401 GTCCCTTCTACTGGAGGGAGTGATGGCAGCACGAGGACAGTTGGAACCCTCCTGCCTCTCATCCCTCCTGGGACAGCTTTCTGGGCAGGTTCGCCTCCTC
110 120 130
LeuGlyAlaLeuGlnGlyLeuLeuGlyThrGlnLeuProLeuGlnGlyArgThrThrAlaHisLysAspProAsnAlaLeuPheLeuSerLeuGlnGlnLeu
501 TTGGGGGCCCTGCAGGGCCTCCTAGGAACCCAGCTTCCTCTACAGGGCAGGACCACAGCTCACAAGGACCCCAATGCCCTCTTCTTGAGCTTGCAACAAC
140 150 160
LeuArgGlyLysValArgPheLeuLeuLeuValGluGlyProThrLeuCysValArgArgThrLeuProThrThrAlaValProSerSerThrSerGln
601 TGCTTCGGGGAAAGGTGCGCTTCCTGCTTCTGGTAGAAGGTCCCACCCTCTGTGTCAGACGGACCCTGCCAACCACAGCTGTCCCAAGCAGTACTTCTCA
170 180 190 200
LeuLeuThrLeuAsnLysPheProAsnArgThrSerGlyLeuLeuGluThrAsnPheSerValThrAlaArgThrAlaGlyProGlyLeuLeuSerArg
701 ACTCCTCACACTAAACAAGTTCCCAAACAGGACTTCTGGATTGTTGGAGACGAACTTCAGTGTCACAGCCAGAACTGCTGGCCCTGGACTTCTGAGCAGG

FIG. 17A

```
                                   210                                                        220                                                         230
     LeuGlnGlyPheArgValLysIleThrProGlyGlnLeuAsnGlnThrSerArgSerProValGlnIleSerGlyTyrLeuAsnArgThrHisGlyProVal
801  CTTCAGGGATTCAGAGTCAAGATTACTCCTGGTCAGCTAAATCAAACCTCCAGGTCCCCAGTCCAAATCTCTGGATACCTGAACAGGACACACGGACCTG

                      240                                    250                                    260
     AsnGlyThrHisGlyLeuPheAlaGlyThrSerLeuGlnThrLeuGluAlaSerAspIleSerProGlyAlaPheAsnLysGlySerLeuAlaPheAsn
901  TGAATGGAACTCATGGGCTCTTTGCTGGAACCTCACTTCAGACCCTGGAAGCCTCAGACATCTCGCCCGGAGCTTTCAACAAAGGCTCCCTGGCATTCAA

           270                            280                                 290                           300
     LeuGlnGlyGlyLeuProProSerProSerLeuAlaProAspGlyHisThrProPheProProSerProAlaLeuProThrThrHisGlySerProPro
1001 CCTCCAGGGTGGACTTCCTCCTTCTCCAAGCCTTGCTCCTGATGGACACACACCCTTCCCTCCTTCACCTGCCTTGCCCACCACCCATGGATCTCCACCC

                                310                                     320                                330
     GlnLeuHisProLeuPheProAspProSerThrThrMetProAsnSerThrAlaProHisProValThrMetTyrProHisProArgAsnLeuSerGlnGlu
1101 CAGCTCCACCCCCTGTTTCCTGACCCTTCCACCACCATGCCTAACTCTACCGCCCCTCATCCAGTCACAATGTACCCTCATCCCAGGAATTTGTCTCAGG

     Thr
1201 AAACATAGcgcgggcactggcccagtgagcgtctgcagcttctctcggggacaagcttccccaggaaggctgagaggcagctgcatctgctccagatgtt

1301 ctgctttcacctaaaggccctggggaagggatacacagcactggagattgtaaaattttaggagctattttttttttaacctatcagcaatattcatcag

1401 agcagctagcgatctttggtctattttcggtataaatttgaaaatcactaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

1501 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

FIG. 17B

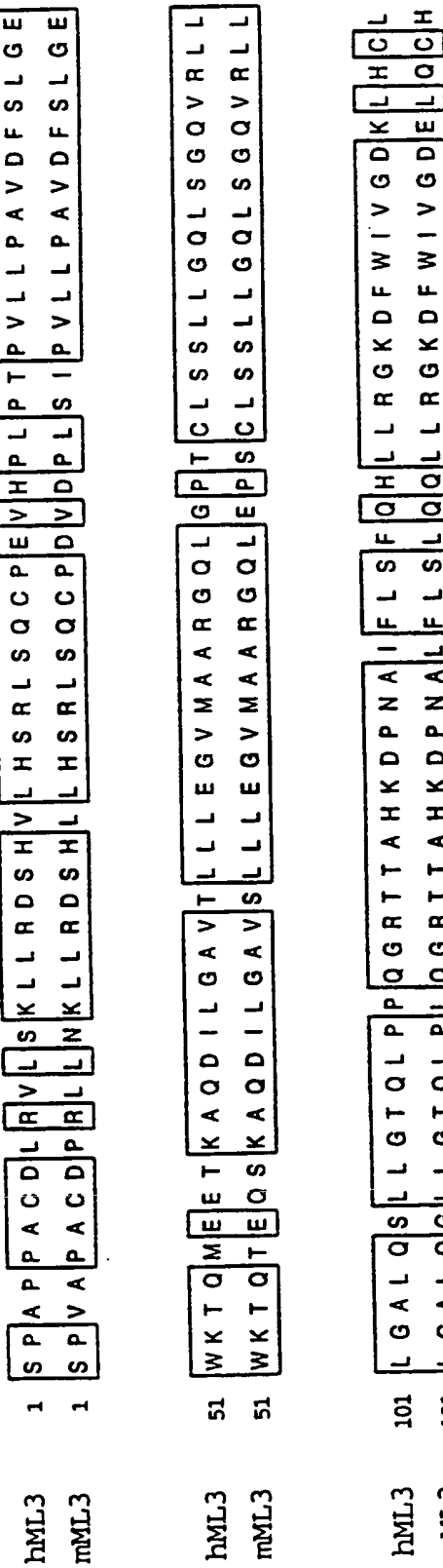

hML3 1 SPAPPACDLRVLSKLLRDSHVLHSRLSQCPEVHPLPTPVLLPAVDFSLGE
mML3 1 SPVAPACDPRLLNKLLRDSHLLHSRLSQCPDVDPLSIPVLLPAVDFSLGE
hML3 51 WKTQMEETKAQDILGAVTLLLEGVMAARGQLGPTCLSSLLGQLSGQVRLL
mML3 51 WKTQTEQSKAQDILGAVSLLLEGVMAARGQLEPSCLSSLLGQLSGQVRLL
hML3 101 LGALQSLLGTQLPPQGRTTAHKDPNAIFLSFQHLLRGKDFWIVGDKLHCL
mML3 101 LGALQGLLGTQLPLQGRTTAHKDPNALFLSLQQLLRGKDFWIVGDELQCH
hML3 151 SQNYWLWASEVAAGIQSQD-SWSAEPNLQVPGPNPRIPEQDTRTLEWNSW
mML3 151 SQNCWPWTSEQASGIQSQDYSWSAKSNLQVPSPNLWIPEQDTRTCEWNSW
hML3 200 TLSWTLTQDPRSPGHFLRNIRHRLPATQPPAWIFSFPNPSSYWTVYALPS
mML3 201 ALCWNLTSDPGSLRHLARSFQQRLPGIQPPGWTSSFSKPCS
hML3 250 STHLAHPCGPAPPPAS

FIG. 18

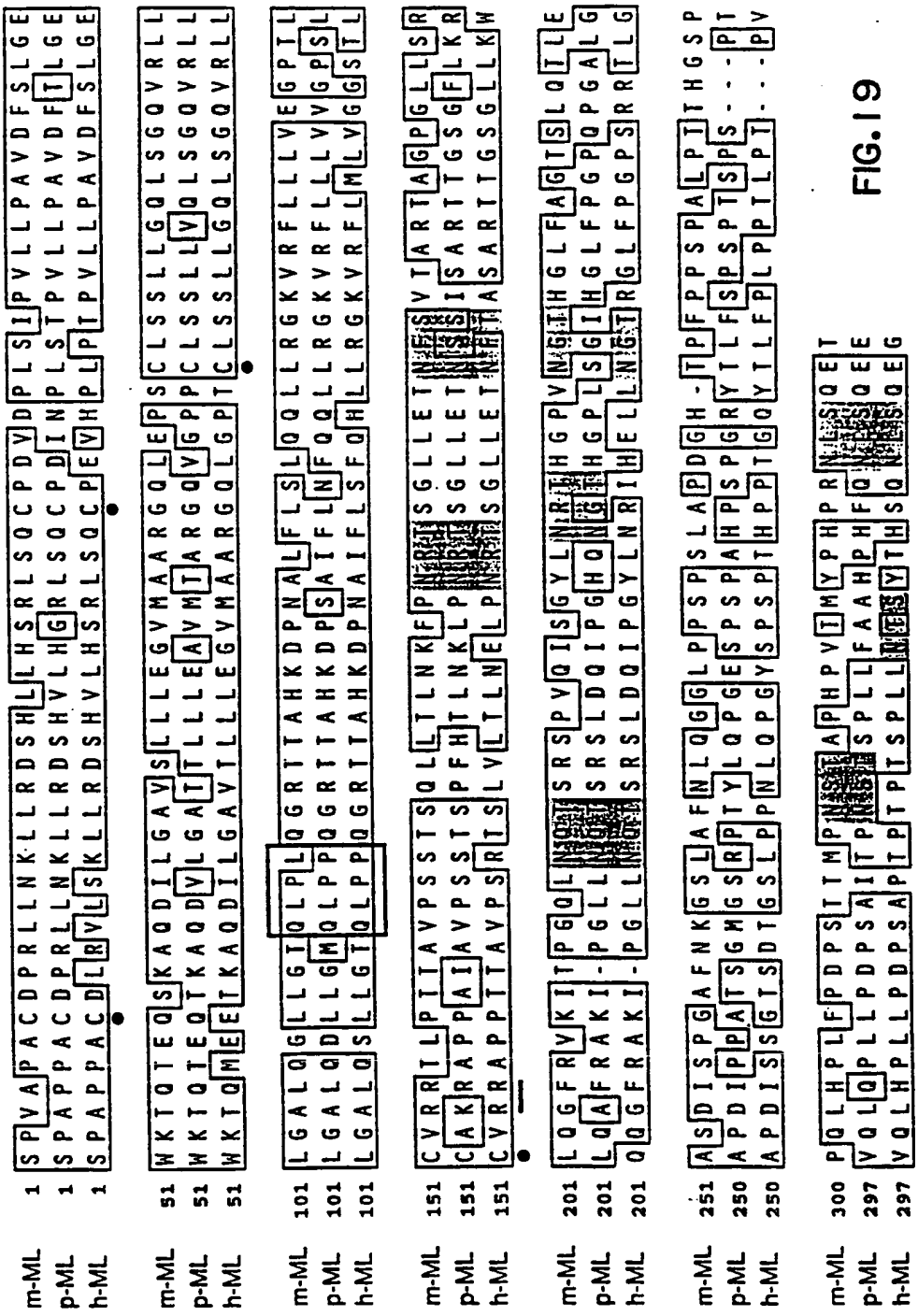


FIG. 19

```
                10                  20                  30
   SerProAlaProProAlaCysAspProArgLeuLeuAsnLysLeuLeuArgAspSerHisValLeuHisGlyArgLeuSerGlnCysProAspIleAsnPro
 1 AGCCCGGCTCCTCCCTGCCTGTGACCCCCGACTCCCTAAATAAACTGCTTCGTGACTCCCATGTCCTTCACGGCAGACTGAGCCAGTGCCCAGACATTAACC

                40                  50                  60
     LeuSerThrProValLeuLeuProAlaValAspPheThrLeuGlyGluTrpLysThrGlnThrGluGlnThrLysAlaGlnAspValLeuGlyAlaThr
101 CTTTGTCCACACCTGTCCTGCTGTGGACTTCACCTTGGGAGAATGGAAAACCCAGACGGAGCAGACAAAGGCACAGGATGTCCTGGGAGCCAC

                70                  80                  90                  100
   ThrLeuLeuLeuGluAlaValMetThrAlaArgGlyGlnValGlyProProCysLeuSerSerLeuLeuValGlnLeuSerGlyGlnValArgLeuLeu
201 AACCCTTCTGCTGGAGGCAGTGATGACAGCACGGGGACAAGTGGGACCCCCTTGCCTCTCATCCCTGCTGGTGCAGCTTTCTGGACAGGTTCGCCTCCTC

                110                 120                 130
   LeuGlyAlaLeuGlnAspLeuLeuGlyMetGlnLeuProProGlnGlyArgThrThrAlaHisLysAspProSerAlaIlePheLeuAsnPheGlnGlnLeu
301 CTCGGGGCCCTGCAGGACCTCCTTGGAATGCAGCTTCCTCCACAGGGAAGGACCACAGCTCACAAGGATCCCAGTGCCATCTTCCTGAACTTCCAACAAC

                140                 150                 160
     LeuArgGlyLysValArgPheLeuLeuValValGlyProSerLeuCysAlaLysArgAlaProProAlaIleAlaValProSerSerThrSerPro
401 TGCTCCGAGGAAAGGTGCGTTTCCTGCTCCTTGTAGTGGGGCCCTCCCTCTGTGCCAAGAGGGCCCCACCCGCCATAGCTGTCCCGAGCAGCACCTCTCC
```

FIG.20A

```
          170                       180                       190                       200
     PheHisThrLeuAsnLysLeuProAsnArgThrSerGlyLeuLeuGluThrAsnSerSerIleSerAlaArgThrThrGlySerGlyPheLeuLysArg
501  ATTCCACACACTGAACAAGCTCCCAAACAGGACCTCTGGATTGTTGGAGACAAACTCCAGTATCTCAGCCAGAACTACTGGCTCTGGATTTCTCAAGAGG

                         210                       220                       230
     LeuGlnAlaPheArgAlaLysIleProGlyLeuLeuAsnGlnThrSerArgSerLeuAspGlnIleProGlyHisGlnAsnGlyThrHisGlyProLeuSer
601  CTGCAGGCATTCAGAGCCAAGATTCCTGGTCTGCTGAACCAAACCTCCAGGTCCCTAGACCAAATCCCTGGACACCAGAATGGGACACACGGACCCTTGA

                    240                       250                       260
      GlyIleHisGlyLeuPheProGlyProGlnProGlyAlaLeuGlyAlaProAspIleProProAlaThrSerGlyMetGlySerArgProThrTyrLeu
701  GTGGAATTCATGGACTCTTTCCTGGACCCCAACCCGGGGCCCTCGGAGCTCCAGACATTCCTCCAGCAACTTCAGGCATGGGCTCCCGGCCAACCTACCT

          270                       280                       290                       300
     GlnProGlyGluSerProSerProAlaHisProSerProGlyArgTyrThrLeuPheSerProSerProThrSerProSerProThrValGlnLeuGln
801  CCAGCCTGGAGAGTCTCCTTCCCCAGCTCACCCTTCTCCTGGACGATACACTCTCTTCTCTCCTTCACCCACCTCGCCCTCCCCCACAGTCCAGCTCCAG

                         310                       320                       330
     ProLeuLeuProAspProSerAlaIleThrProAsnSerThrSerProLeuLeuPheAlaAlaHisProHisPheGlnAsnLeuSerGlnGluGlu
901  CCTCTGCTTCCTGACCCCTCTGCGATCACACCCAACTCTACCAGTCCTCTTCTATTTGCAGCTCACCCTCATTTCCAGAACCTGTCTCAGGAAGAGTAAG
```

FIG.20B

```
1001 GTGCTCAGACCCTGCCAACTTCAGCA
```

```
                    10                  20                  30
   SerProAlaProProAlaCysAspProArgLeuLeuAsnLysLeuLeuArgAspSerHisValLeuHisGlyArgLeuSerGlnCysProAspIleAsnPro
 1 AGCCCGGCTCCCTCCTGCCTGTGACCCCCGACTCCCTAAATAAACTGCTTCGTGACTCCCCATGTCCTTCACGGCAGACTGAGCCAGTGCCCAGACATTAACC

                    40                  50                  60
     LeuSerThrProValLeuLeuProAlaValAspPheThrLeuGlyGluTrpLysThrGlnThrGluGlnThrLysAlaGlnAspValLeuGlyAlaThr
101 CTTTGTCCACACCTGTCCTGCCTGCTGTGGACTTCACCTTGGGAGAATGGAAAACCCAGACGGAGGCAGACAAAGGCACAGGATGTCCTGGGAGCCAC

                    70                  80                  90                  100
     ThrLeuLeuLeuGluAlaValMetThrAlaArgGlyGlnValGlyProProCysLeuSerSerLeuLeuValGlnLeuSerGlyGlnValArgLeuLeu
201 AACCCTTCTGCTGGAGGCAGTGATGACAGCACGGGGACAAGTGGGACCCCCTTGCCTCTCATCCCTGCTGGTGCAGCTTTCTGGACAGGTTCGCCTCCTC

                    110                 120                 130
     LeuGlyAlaLeuGlnAspLeuLeuGlyMetGlnGlyArgThrThrAlaHisLysAspProSerAlaIlePheLeuAsnPheGlnGlnLeuLeuArgGlyLys
301 CTCGGGGCCCTGCAGGACCTCCTTGGAATGCAGGGAAGGACCACAGCTCACAAGGATCCCAGTGCCATCTTCCTGAACTTCCAACAACTGCTCCGAGGAA

                    140                 150                 160
     ValArgPheLeuLeuLeuValValGlyProSerLeuCysAlaLysArgAlaProProAlaIleAlaValProSerSerSerThrSerProPheHisThrLeu
401 AGGTGCGTTTCCTGCTCCTTGTAGTGGGGCCCTCCCTGTGCCAAGAGGGCCCCACCCGCCATAGCTGTCCCGAGCAGCACCTCTCCATTCCACACACT
```

FIG.21A

```
          170                                 180                                 190                                 200
      AsnLysLeuProAsnArgThrSerGlyLeuLeuGluThrAsnSerSerIleSerAlaArgThrThrGlySerGlyPheLeuLysArgLeuGlnAlaPhe
501   GAACAAGCTCCCAAACAGGACCTCTGGATTGTTGGAGACAAACTCCAGTATCTCAGCCAGAACTACTGGCTCTGGATTTCTCAAGAGGCTGCAGGCATTC

                          210                                 220                                 230
      ArgAlaLysIleProGlyLeuLeuAsnGlnThrSerArgSerLeuAspGlnIleProGlyHisGlnAsnGlyThrHisGlyProLeuSerGlyIleHisGly
601   AGAGCCAAGATTCCTGGTCTGCTGAACCAAACCTCCAGGTCCCTAGACCAAATCCCTGGACACCAGAATGGGACACACGGACCCTTGAGTGGAATTCATG

                    240                                 250                                 260
      LeuPheProGlyProGlnProGlyAlaLeuGlyAlaProAspIleProProAlaThrSerGlyMetGlySerArgProThrTyrLeuGlnProGlyGlu
701   GACTCTTTCCTGGACCCCAACCCGGGGCCCTCGGAGCTCCAGACATTCCTCCAGCAACTTCAGGCATGGGCTCCCGGCCAACCTACCTCCAGCCTGGAGA

          270                                 280                                 290                                 300
      SerProSerProAlaHisProSerProGlyArgTyrThrLeuPheSerProSerProThrSerProSerProThrValGlnLeuGlnProLeuLeuPro
801   GTCTCCTTCCCCAGCTCACCCTTCTCCTGGACGATACACTCTCTTCTCTCCTTCACCCACCTCGCCCTCCCCCACAGTCCAGCTCCAGCCTCTGCTTCCT

                          310                                 320
      AspProSerAlaIleThrProAsnSerThrSerProLeuLeuPheAlaAlaHisProHisPheGlnAsnLeuSerGlnGluGlu
901   GACCCCTCTGCGATCACACCCAACTCTACCAGTCCTCTTCTATTTGCAGCTCACCCTCATTTCCAGAACCTGTCTCAGGAAGAGTAAGGTGCTCAGACCC
```

FIG.21B

```
1001  TGCCAACTTCAGCA
```

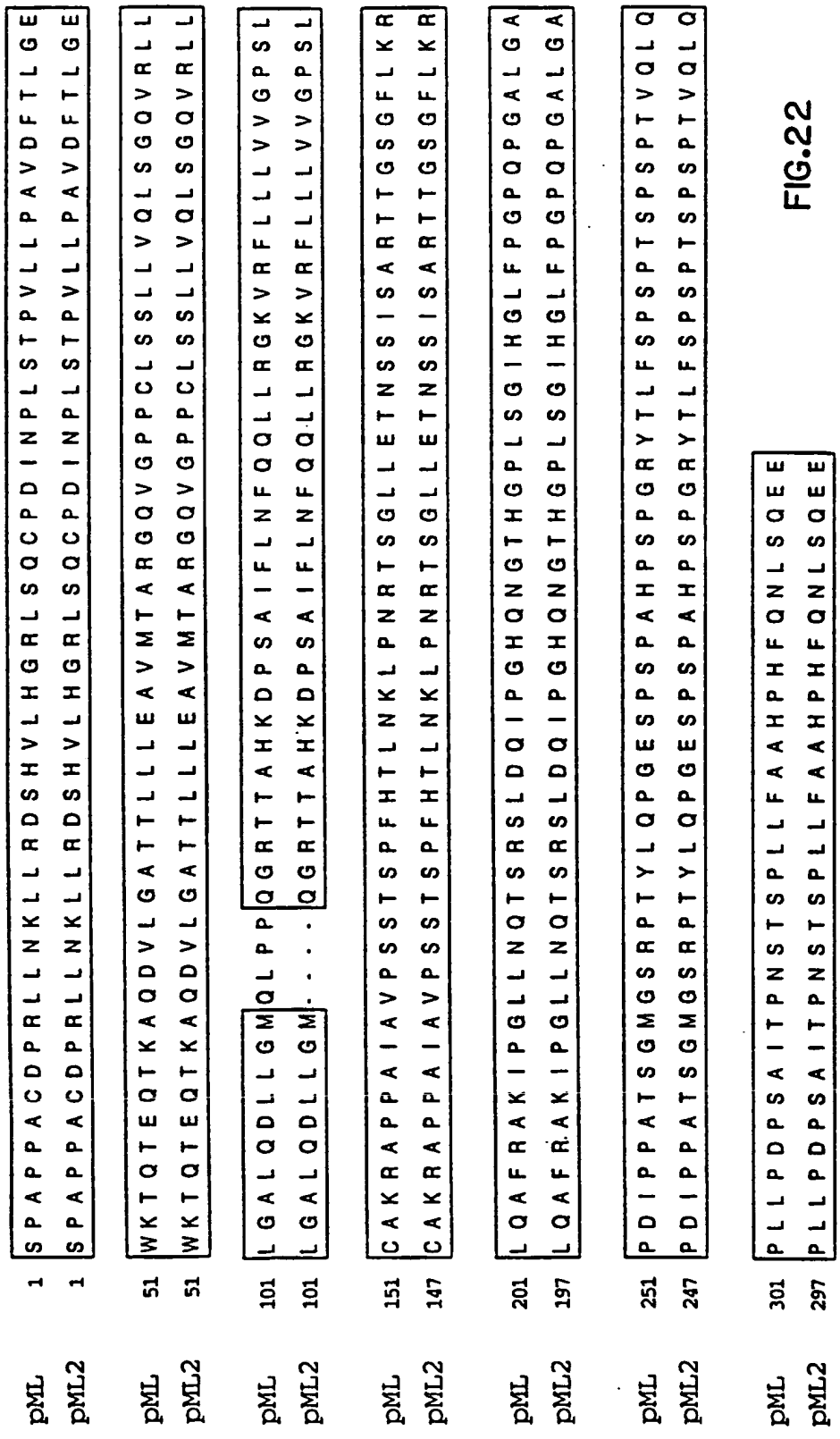

pML 1 SPAPPACDPRLLNKLLRDSHVLHGRLSQCPDINPLSTPVLLPAVDFTLGE
pML2 1 SPAPPACDPRLLNKLLRDSHVLHGRLSQCPDINPLSTPVLLPAVDFTLGE
pML 51 WKTQTEQTKAQDVLGATTLLLEAVMTARGQVGPPCLSSLLVQLSGQVRLL
pML2 51 WKTQTEQTKAQDVLGATTLLLEAVMTARGQVGPPCLSSLLVQLSGQVRLL
pML 101 LGALQDLLGMQLPPQGRTTAHKDPSAIFLNFQQLLRGKVRFLLLVVGPSL
pML2 101 LGALQDLLGM. . . QGRTTAHKDPSAIFLNFQQLLRGKVRFLLLVVGPSL
pML 151 CAKRAPPAIAVPSSTSPFHTLNKLPNRTSGLLETNSSISARTTGSGFLKR
pML2 147 CAKRAPPAIAVPSSTSPFHTLNKLPNRTSGLLETNSSISARTTGSGFLKR
pML 201 LQAFRAKIPGLLNQTSRSLDQIPGHQNGTHGPLSGIHGLFPGPQPGALGA
pML2 197 LQAFRAKIPGLLNQTSRSLDQIPGHQNGTHGPLSGIHGLFPGPQPGALGA
pML 251 PDIPPATSGMGSRPTYLQPGESPSPAHPSPGRYTLFSPSPSPTVQLQ
pML2 247 PDIPPATSGMGSRPTYLQPGESPSPAHPSPGRYTLFSPSPSPTVQLQ
pML 301 PLLPDPSAITPNSTSPLLFAAHPHFQNLSQEE
pML2 297 PLLPDPSAITPNSTSPLLFAAHPHFQNLSQEE

FIG.22

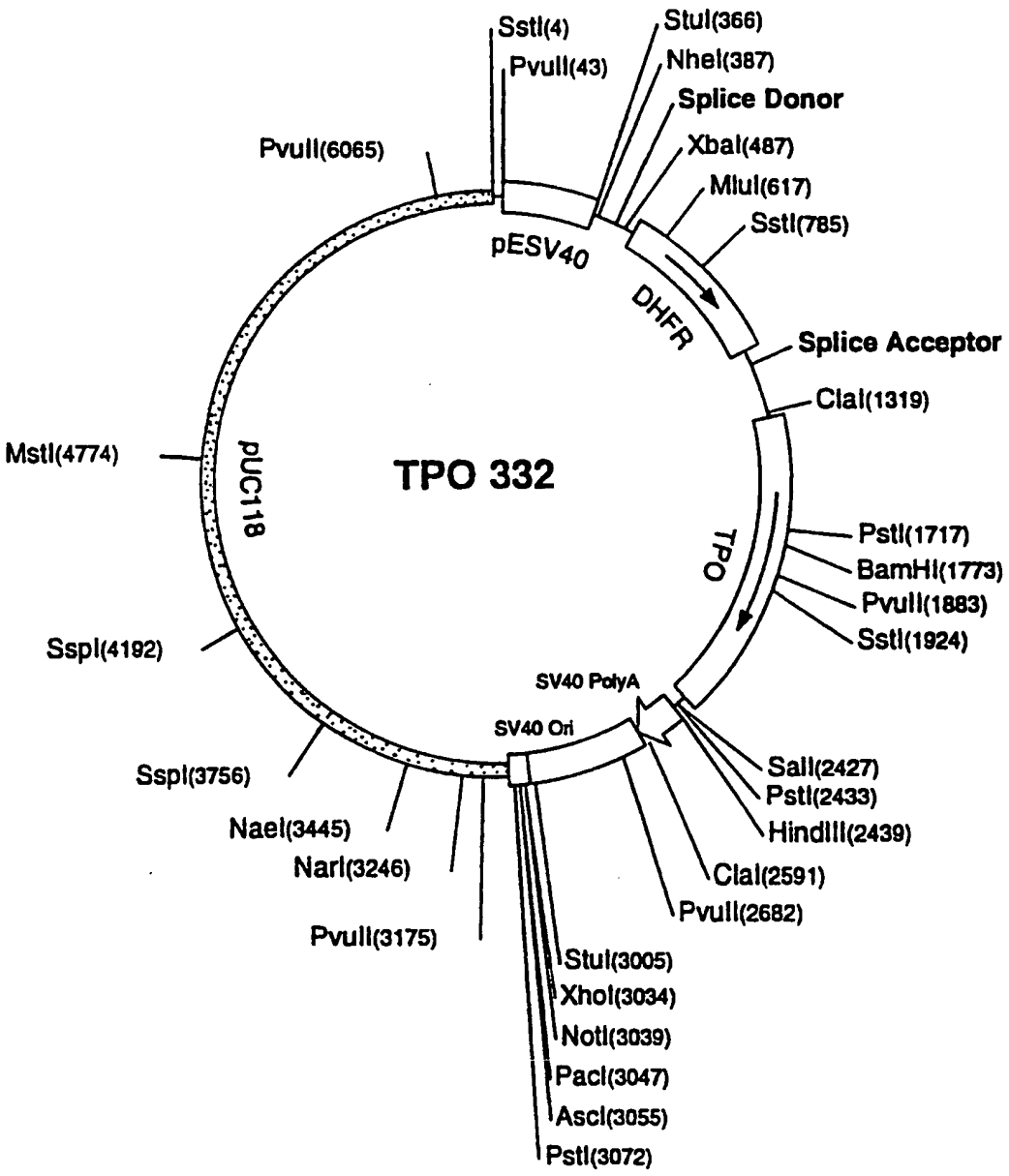

SstI(4)
PvuII(43)
StuI(366)
NheI(387)
Splice Donor
XbaI(487)
MluI(617)
SstI(785)
Splice Acceptor
ClaI(1319)
PstI(1717)
BamHI(1773)
PvuII(1883)
SstI(1924)
SalI(2427)
PstI(2433)
HindIII(2439)
ClaI(2591)
PvuII(2682)
StuI(3005)
XhoI(3034)
NotI(3039)
PacI(3047)
AscI(3055)
PstI(3072)
PvuII(3175)
NarI(3246)
NaeI(3445)
SspI(3756)
SspI(4192)
MstI(4774)
PvuII(6065)
pESV40
DHFR
TPO
SV40 PolyA
SV40 Ori
pUC118
TPO 332

FIG.23

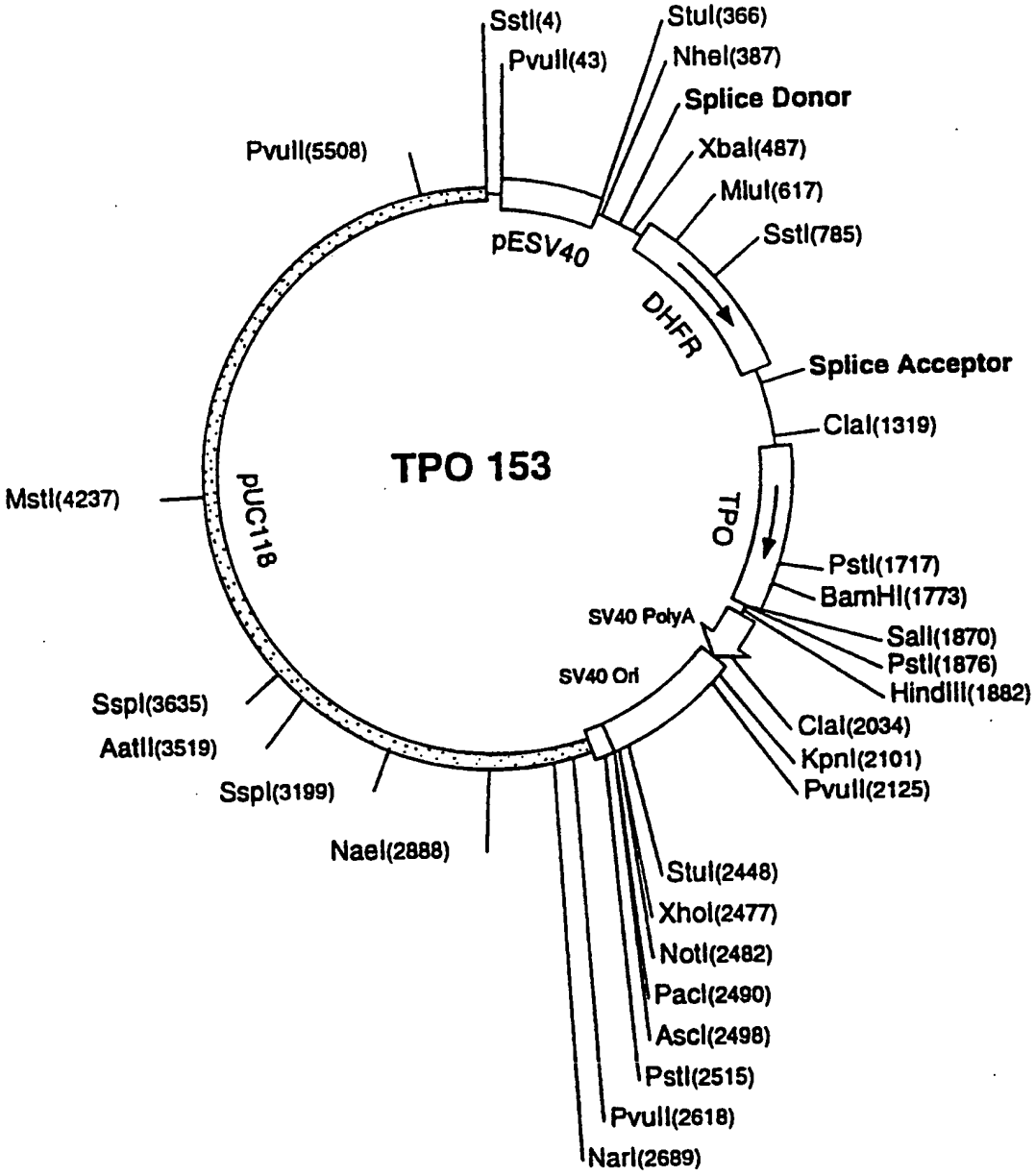

SstI(4)
PvuII(43)
StuI(366)
NheI(387)
Splice Donor
XbaI(487)
MluI(617)
SstI(785)
Splice Acceptor
ClaI(1319)
PstI(1717)
BamHI(1773)
SalI(1870)
PstI(1876)
HindIII(1882)
ClaI(2034)
KpnI(2101)
PvuII(2125)
StuI(2448)
XhoI(2477)
NotI(2482)
PacI(2490)
AscI(2498)
PstI(2515)
PvuII(2618)
NarI(2689)
NaeI(2888)
SspI(3199)
AatII(3519)
SspI(3635)
MstI(4237)
PvuII(5508)
pESV40
DHFR
TPO
SV40 PolyA
SV40 Ori
pUC118
TPO 153

FIG.24

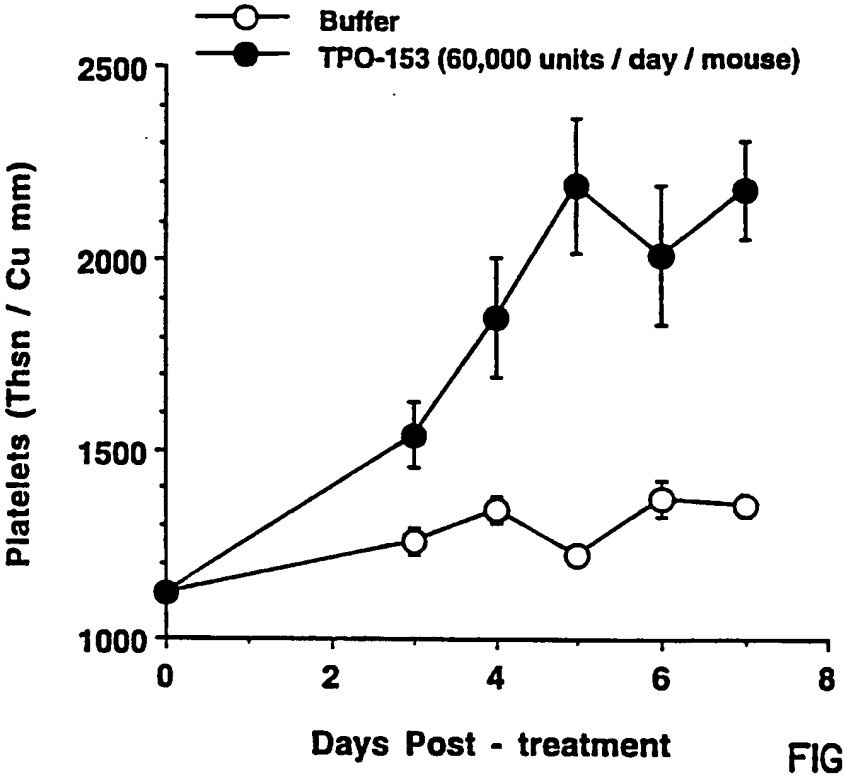


FIG.25A

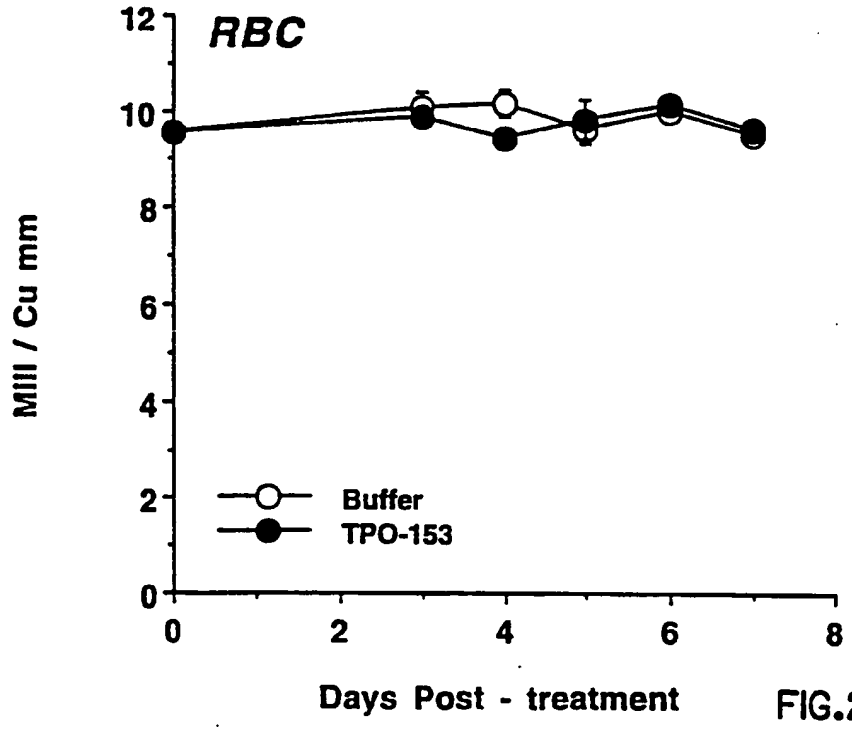


FIG.25B

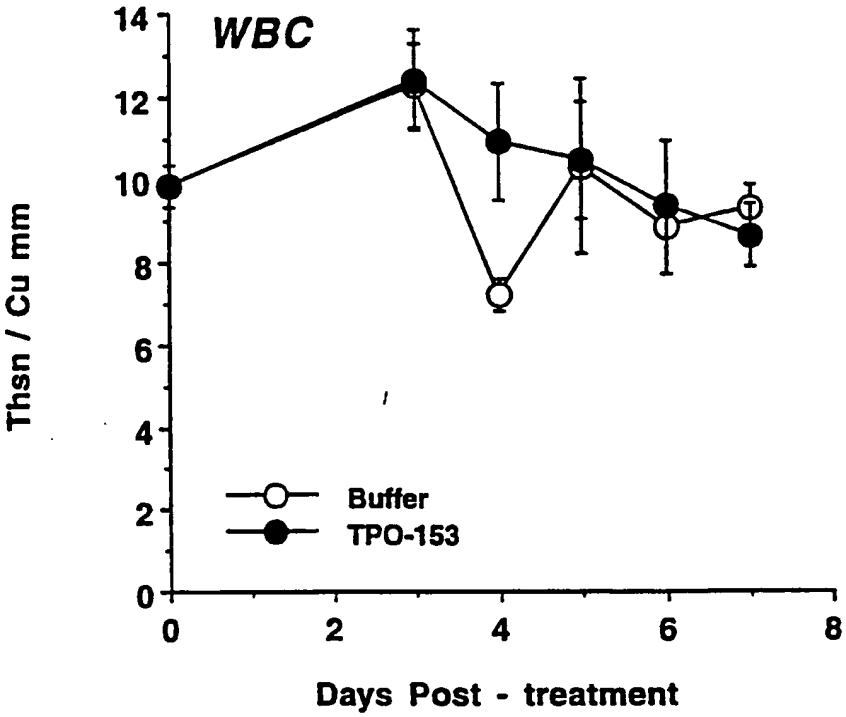


FIG.25C

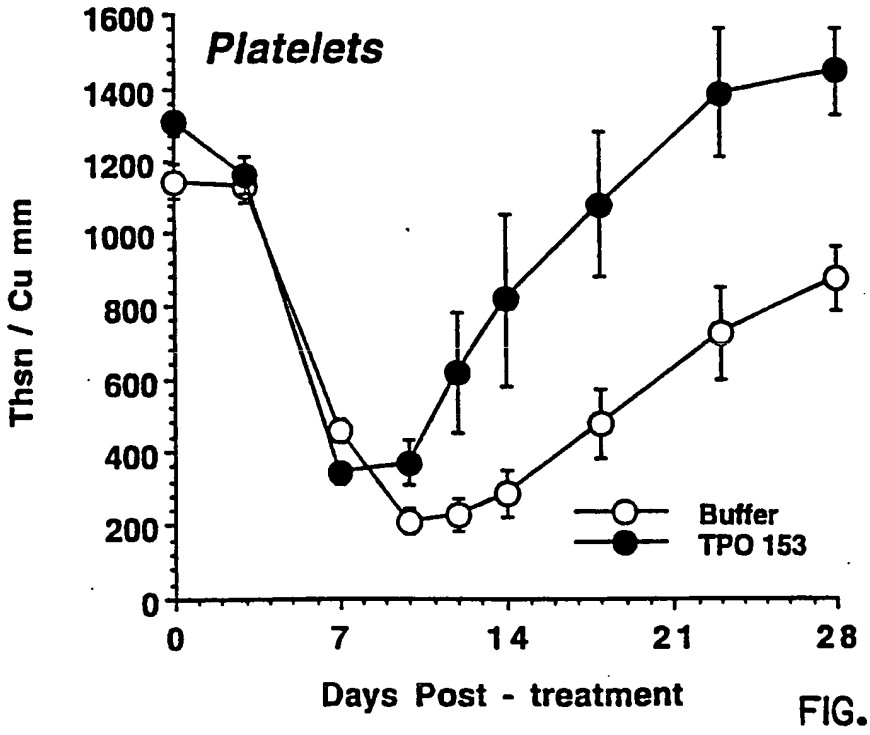


FIG.26A

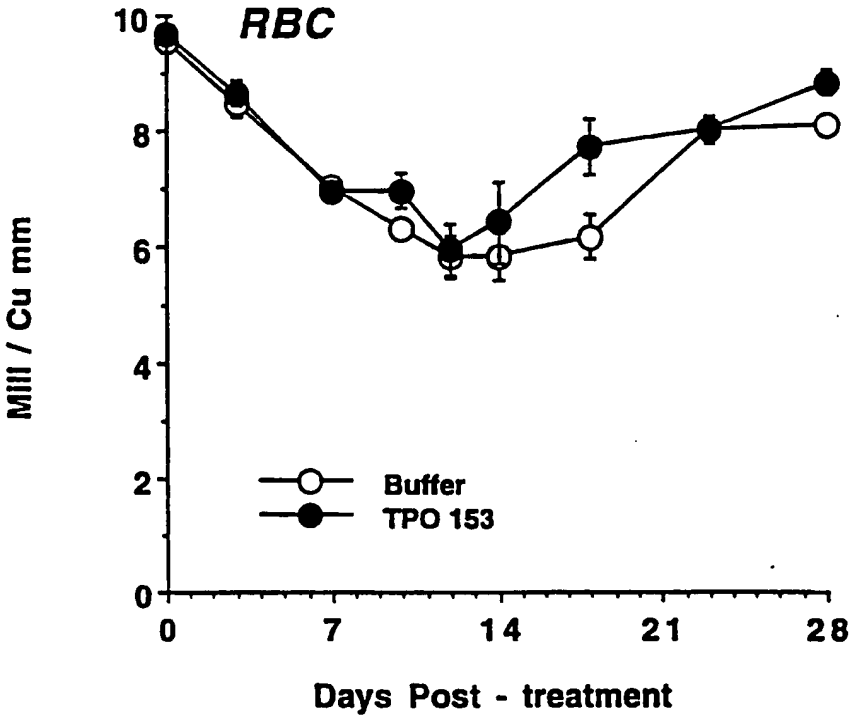


FIG.26B

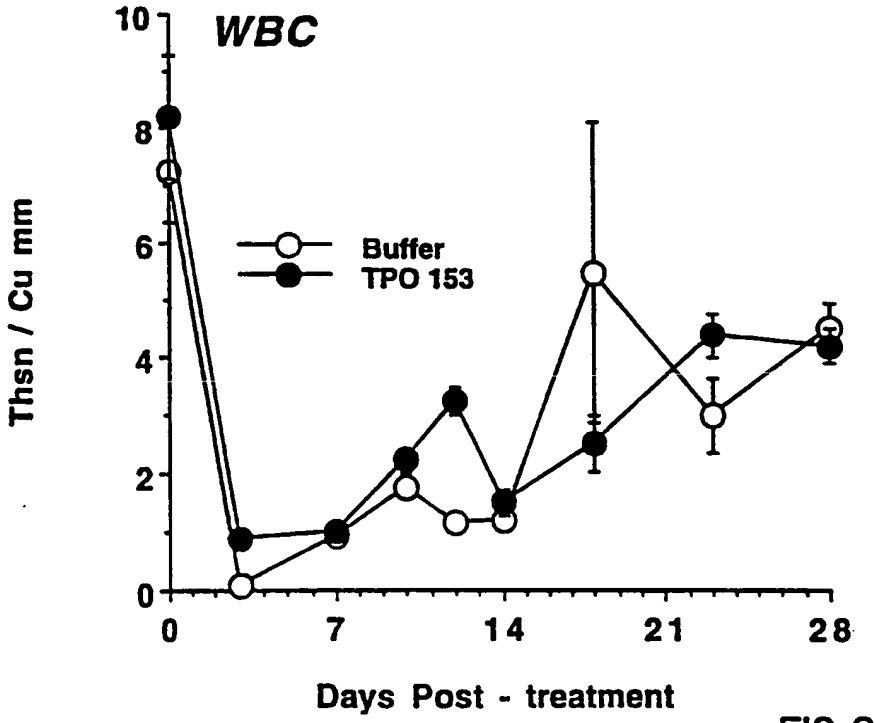


FIG.26C

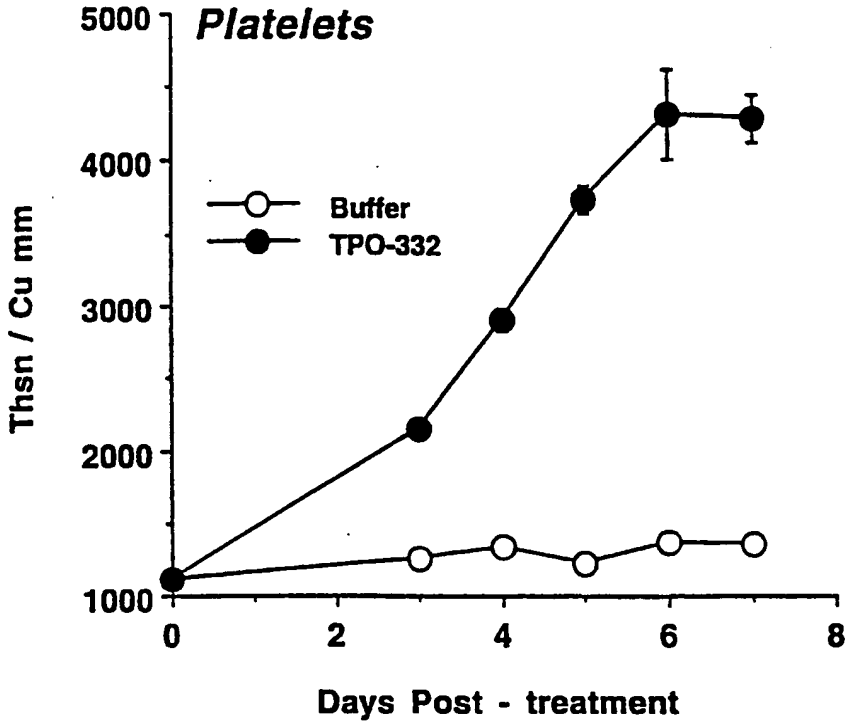


FIG.27A

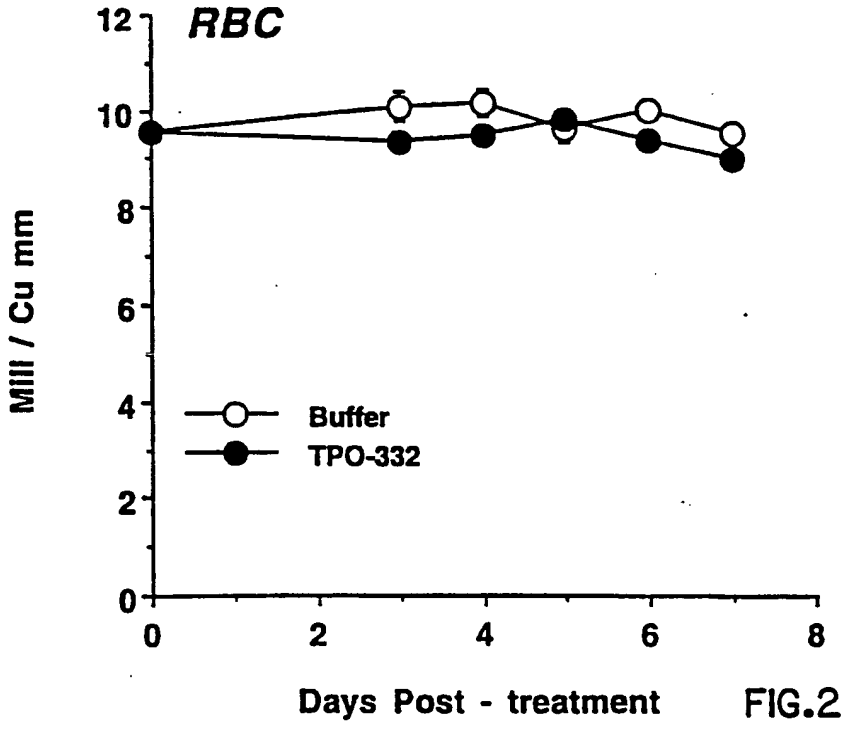


FIG.27B

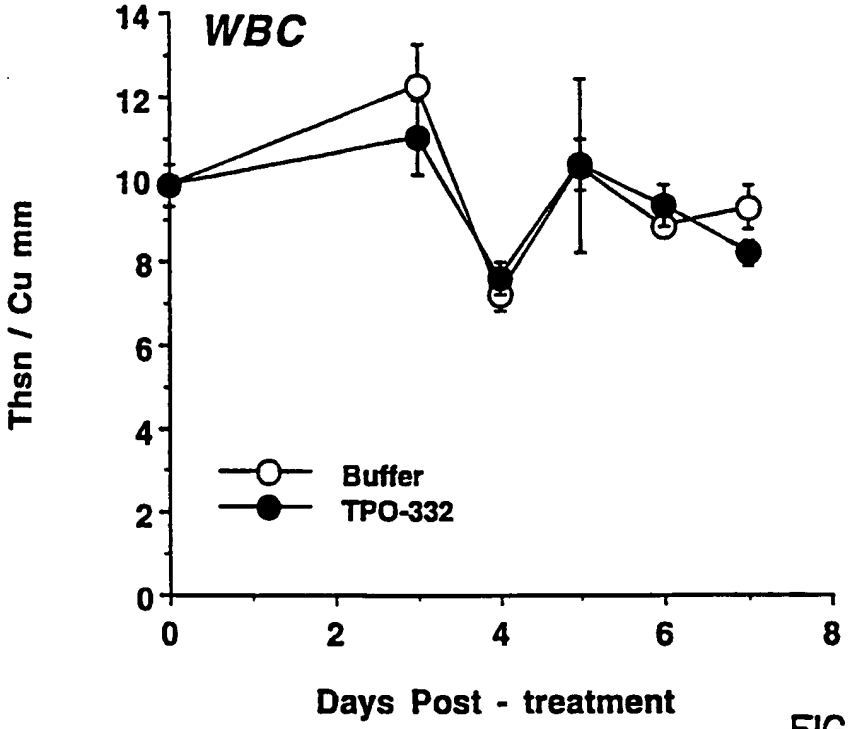


FIG.27C

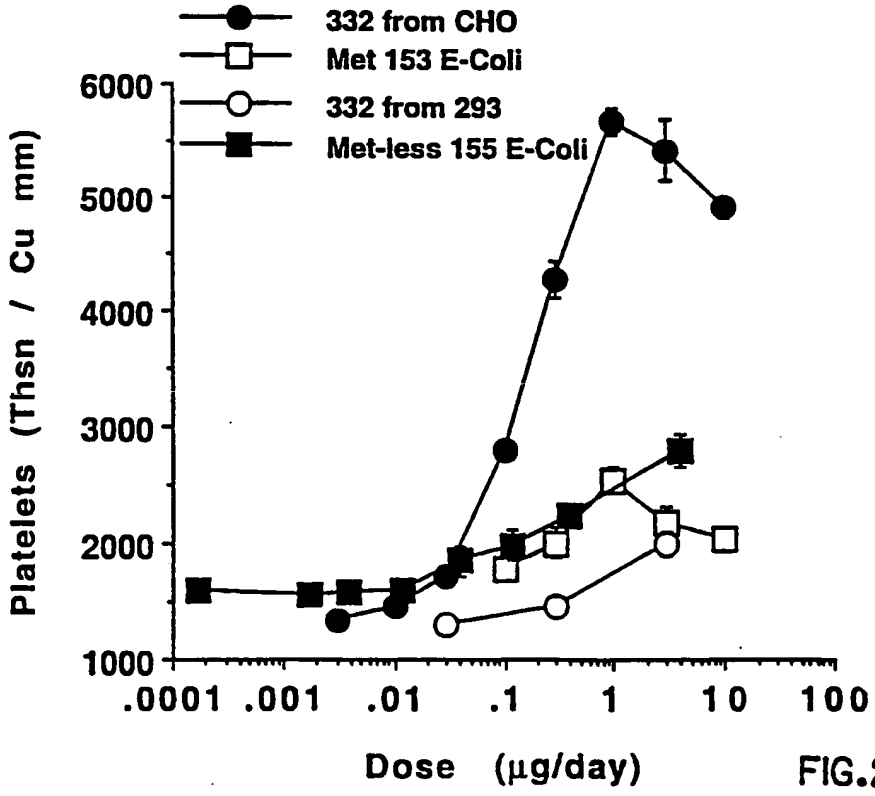


FIG.28

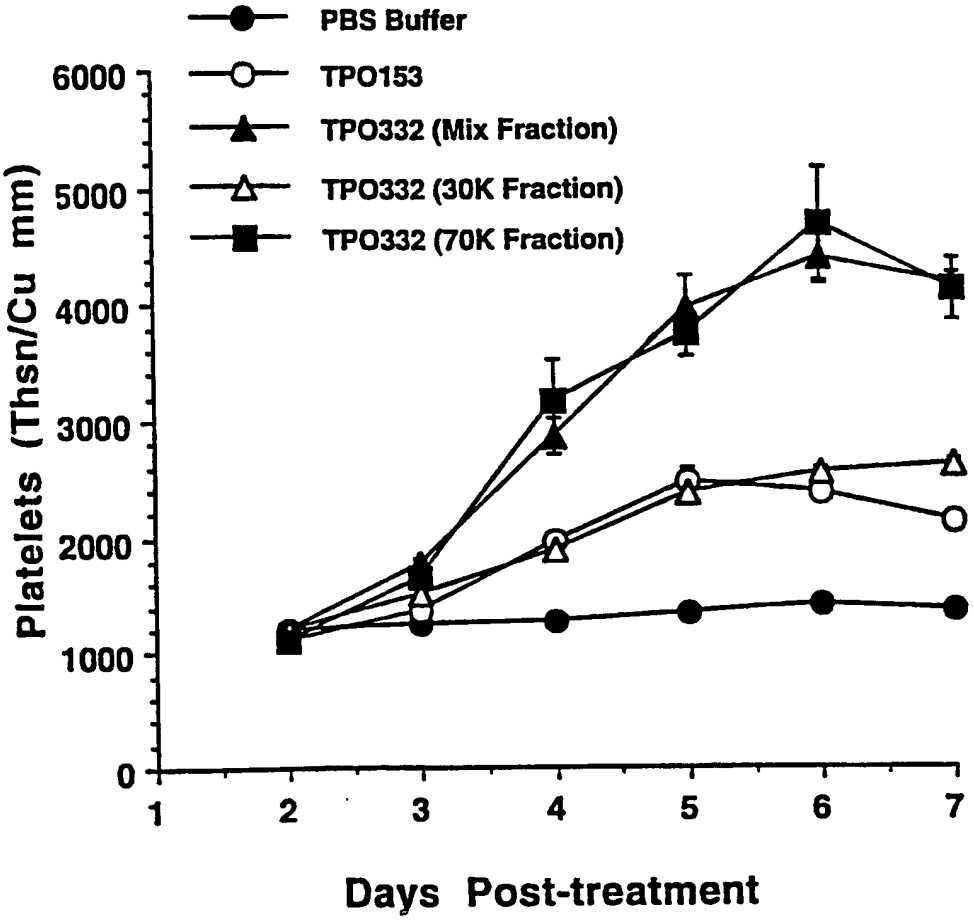

PBS Buffer
TPO153
TPO332 (Mix Fraction)
TPO332 (30K Fraction)
TPO332 (70K Fraction)
Platelets (Thsn/Cu mm)
6000
5000
4000
3000
2000
1000
0
1
2
3
4
5
6
7
Days Post-treatment

FIG.29

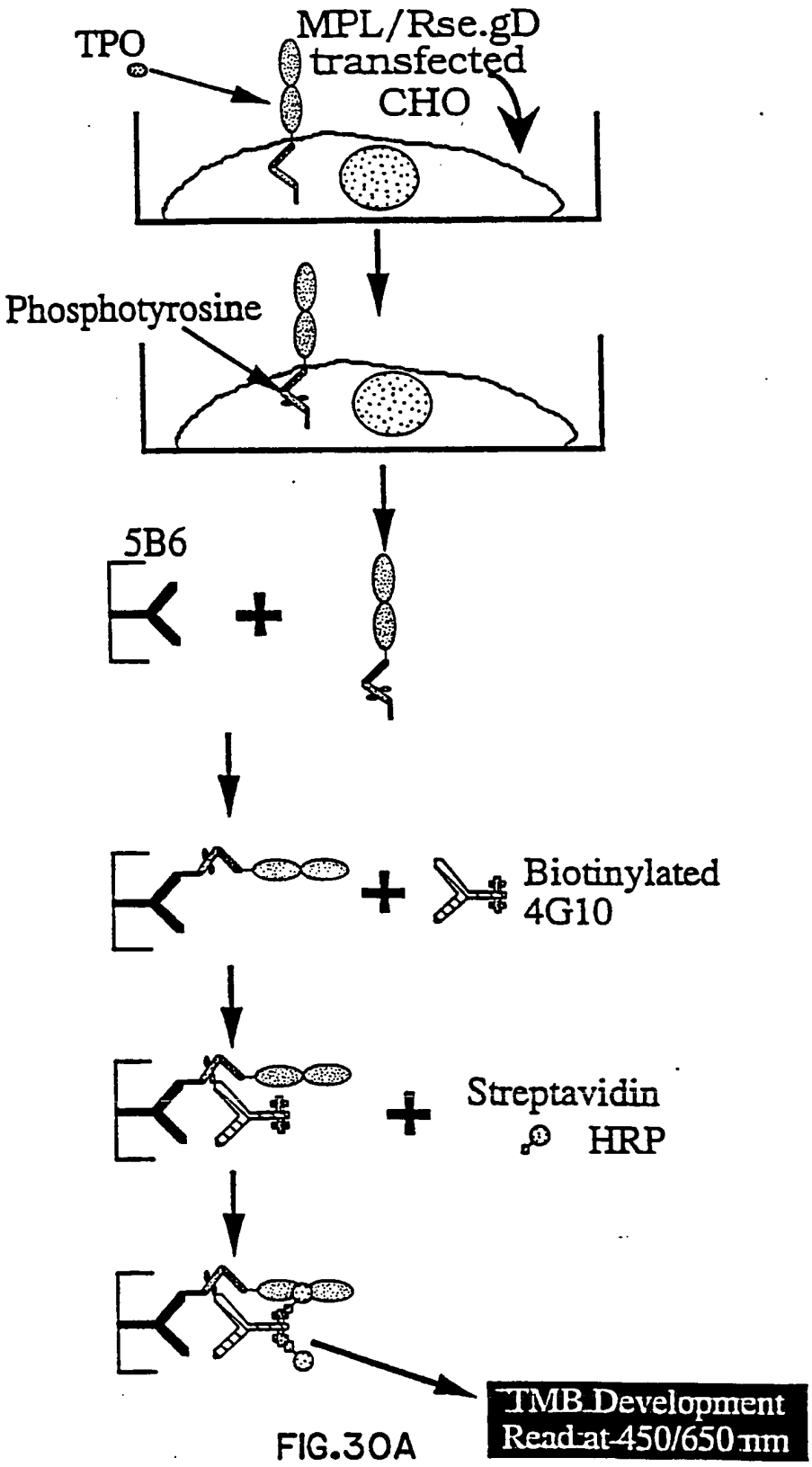
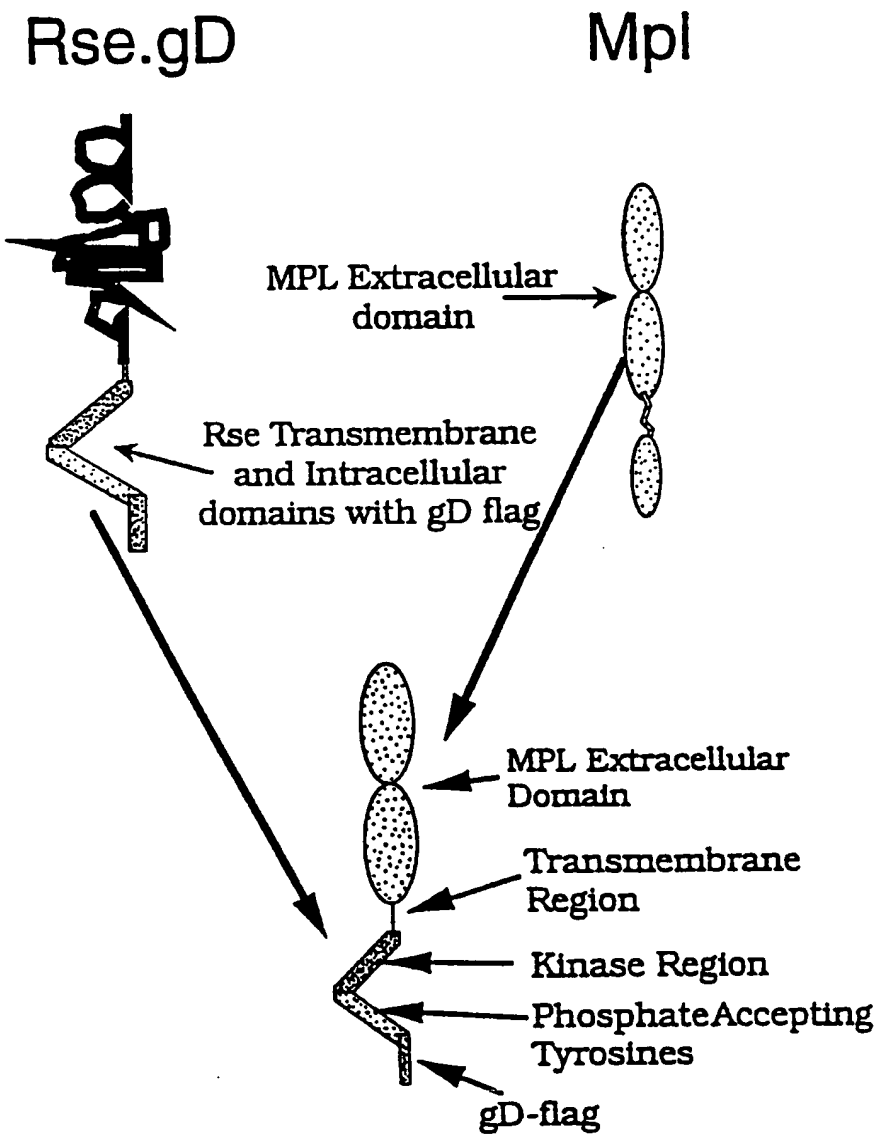

TPO
MPL/Rse.gD transfected CHO
Phosphotyrosine
5B6
Biotinylated 4G10
Streptavidin
HRP
TMB Development
Read at 450/650 nm

FIG.30A



MPL/Rse.gD

FIG.30B

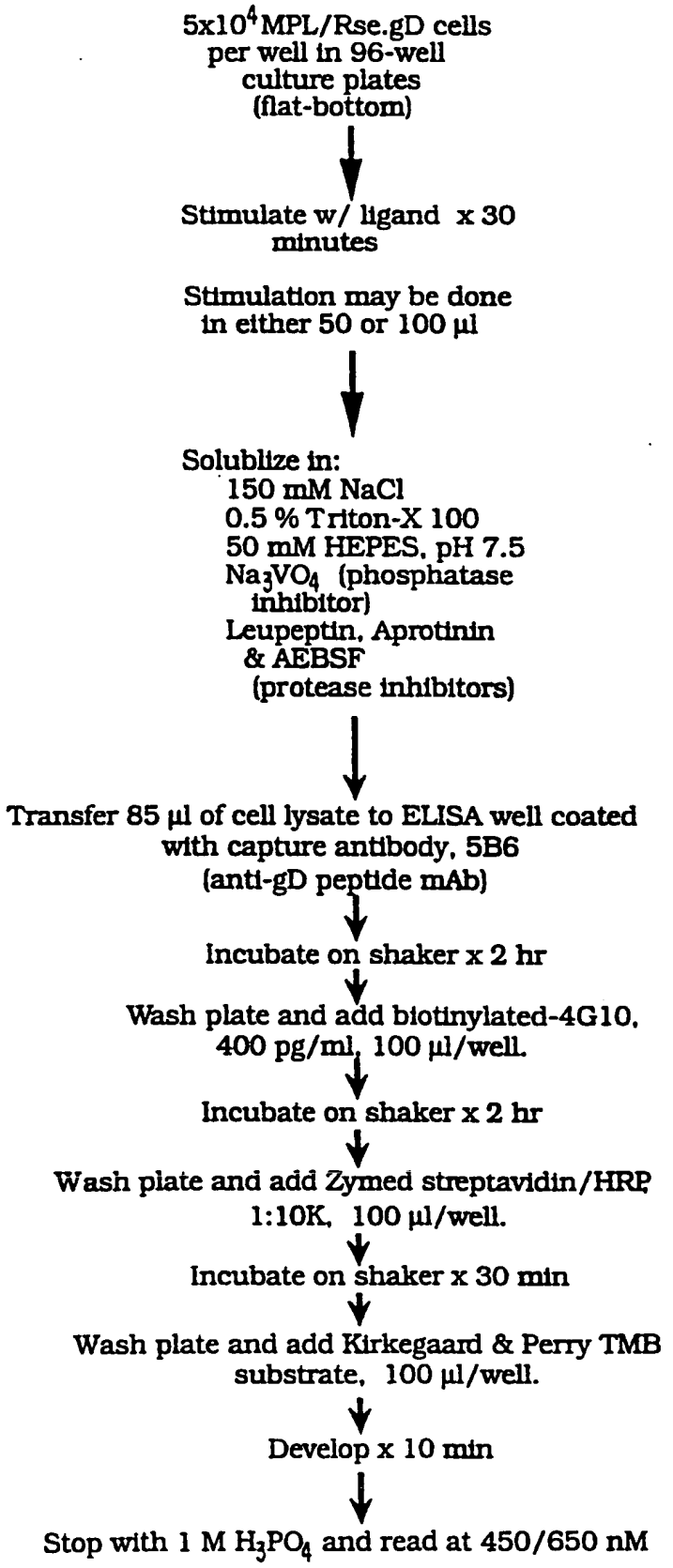

5x10⁴MPL/Rse.gD cells per well in 96-well culture plates (flat-bottom)
Stimulate w/ ligand x 30 minutes
Stimulation may be done in either 50 or 100 µl
Solublize in:
150 mM NaCl
0.5 % Triton-X 100
50 mM HEPES, pH 7.5
Na3VO4 (phosphatase inhibitor)
Leupeptin, Aprotinin & AEBSF (protease inhibitors)
Transfer 85 µl of cell lysate to ELISA well coated with capture antibody, 5B6 (anti-gD peptide mAb)
Incubate on shaker x 2 hr
Wash plate and add biotinylated-4G10, 400 pg/ml, 100 µl/well.
Incubate on shaker x 2 hr
Wash plate and add Zymed streptavidin/HRP 1:10K, 100 µl/well.
Incubate on shaker x 30 min
Wash plate and add Kirkegaard & Perry TMB substrate, 100 µl/well.
Develop x 10 min
Stop with 1 M H3PO4 and read at 450/650 nM

FIG.31

```
                 aluI
                sstI
                sacI
                hgiJII
                hgiAI/aspHI
                ecl136II
                bsp1286
                bsiHKAI
                bmyI
                banII
              taqI
   1 TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT CAGTTAGGGT
     AAGCTCGAGC GGGCTGTAAC TAATAACTGA TCTCAGCTAG CTGTCGACAC CTTACACACA GTCAATCCCA
```

```
                                                            aluI
                                                    sau3AI pvuII
                                                    mboI/ndeII[dam-]
                                                    dpnI[dam+]
                                                   pvuI/bspCI
                                               pleI dpnII[dam-]
                                               hinfI  taqI[dam-]
                                           rmaI    mcrI    nspBII
                                           maeI   taqI[dam-]
```

```
                  nlaIV                                      sfaNI                  scrFI
                scrFI                                      ppu10I                   mvaI
                mvaI                                       nsiI/avaIII              ecoRII
                ecoRII                                    nlaIII                    dsaV
                dsaV                                     sphI                       bstNI
                bstNI                                    nspI                       apyI
                apyI[dcm+]                               nspHI                     sexAI
               bsaJI
  71 GTGGAAAGTC CCCAGGCTCC CCAGCAGGCA GAAGTATGCA AAGCATGCAT CTCAATTAGT CAGCAACCAG
     CACCTTTCAG GGGTCCGAGG GGTCGTCCGT CTTCATACGT TTCGTACGTA GAGTTAATCA GTCGTTGGTC
```

FIG.32A

```
                          nlaIV
                 scrFI                                         sfaNI
                       mvaI                                              ppu10I
                 ecoRII                                      nsiI/avaIII
                 dsaV                                       nlaIII
                       bstNI                                         sphI
     [dcm+]            apyI[dcm+]                                    nspI
                      bsaJI                                          nspHI
141 GTGTGGAAAG TCCCCAGGCT CCCCAGCAGG CAGAAGTATG CAAAGCATGC ATCTCAATTA GTCAGCAACC
    CACACCTTTC AGGGGTCCGA GGGGTCGTCC GTCTTCATAC GTTTCGTACG TAGAGTTAAT CAGTCGTTGG

                                                                            nlaIII
                                                                           styI
                                                                           ncoI
                                  aciI                                bslI dsaI
          aciI          aciI  fokI                  aciI bsrI  aciI   aciI bsaJI
211 ATAGTCCCGC CCCTAACTCC GCCCATCCCG CCCCTAACTC CGCCCAGTTC CGCCCATTCT CCGCCCCATG
    TATCAGGGCG GGGATTGAGG CGGGTAGGGC GGGGATTGAG GCGGGTCAAG GCGGGTAAGA GGCGGGGTAC

                                              fnu4HI
                                              bglI
                                             sfiI
                                             haeIII/palI
                                           mnlI   mnlI       ddeI
                                      haeIII/palI bsaJI  mnlI   aluI
                                    mnlI bsaJI aciI    haeIII/palI               mnlI
281 GCTGACTAAT TTTTTTTATT TATGCAGAGG CCGAGGCCGC CTCGGCCTCT GAGCTATTCC AGAAGTAGTG
    CGACTGATTA AAAAAAATAA ATACGTCTCC GGCTCCGGCG GAGCCGGAGA CTCGATAAGG TCTTCATCAC
```

FIG.32B

```
                      rmaI
                     styI
                     bsaJI
                     blnI
                     avrII                      aluI      haeIII/palI
                   haeIII/palI                 rmaI       mcrI
                  stuI                         maeI       eagI/xmaIII/eclXI
    mnlI          haeI                        nheI        eaeI
                 mnlI maeI                   aluI         cfrI
                                                         mspI
                                                         hpaII
351 AGGAGGCTTT TTTGGAGGCC TAGGCTTTTG CAAAAAGCTA GCTTATCCGG
    TCCTCCGAAA AAACCTCCGG ATCCGAAAAC GTTTTTCGAT CGAATAGGCC

                              tfiI
    scrFI                     hinFI
    nciI                    aciI
    mspI                   thaI
    hpaII                  fnuDII/mvnI                        aciI
    dsaV                   bstUI                      maeII  rsaI
    cauII                  bsh1236I                maeIII    csp6I  scfI
401 CCGGGAACGG TGCATTGGAA CGCGGATTCC CCGTGCCAAG AGTGACGTAA GTACCGCCTA TAGAGCGATA
    GGCCCTTGCC ACGTAACCTT GCGCCTAAGG GGCACGGTTC TCACTGCATT CATGGCGGAT ATCTCGCTAT
                                                  ^splice donor

                  fnu4HI
                  bbvI
                 nspBII                                                pflMI
     mnlI        aciI        nlaIII taqI          sfaNI                bslI
471 AGAGGATTTT ATCCCCGCTG CCATCATGGT TCGACCATTG AACTGCATCG TCGCCGTGTC CCAAAATATG
    TCTCCTAAAA TAGGGGCGAC GGTAGTACCA AGCTGGTAAC TTGACGTAGC AGCGGCACAG GGTTTTATAC
                   DHFR ATG^
```

FIG.32C

```
                                    haeIII/palI
                                   haeI
                                  scrFI
                                  mvaI      bsrBI
                                  ecoRII
                                  dsaV
                                  bstNI     aciI                  rsaI
                       bsmAI      apyI[dcm+]          xmnI        csp6I
                       bsaI     bsaJI  mnlI  ddeI   asp700      scaI
541 GGGATTGGCA AGAACGGAGA CCTACCCTGG CCTCCGCTCA GGAACGAGTT CAAGTACTTC CAAAGAATGA
    CCCTAACCGT TCTTGCCTCT GGATGGGACC GGAGGCGAGT CCTTGCTCAA GTTCATGAAG GTTTCTTACT

                                                                  scrFI
                                                                  mvaI
                                                                  ecoRII
            eco57I                                                dsaV
           mboII                      tfiI                        bstNI
          earI/ksp632I                hinfI                       apyI[dcm+]
         mnlI                       alwNI    hphI                 sexAI            ddeI
611 CCACAACCTC TTCAGTGGAA GGTAAACAGA ATCTGGTGAT TATGGGTAGG AAAACCTGGT TCTCCATTCC
    GGTGTTGGAG AAGTCACCTT CCATTTGTCT TAGACCACTA ATACCCATCC TTTTGGACCA AGAGGTAAGG

          tfiI       tru9I          tru9I
          hinfI      mseI           mseI          ddeI
      mboII taqI   ahaIII/draI   aseI/asnI/vspI                      bslI   mnlI
681 TGAGAAGAAT CGACCTTTAA AGGACAGAAT TAATATAGTT CTCAGTAGAG AACTCAAAGA ACCACCACGA
    ACTCTTCTTA GCTGGAAATT TCCTGTCTTA ATTATATCAA GAGTCATCTC TTGAGTTTCT TGGTGGTGCT
```

FIG.32D

```
    sstI
    sacI
    hgiJII
    hgiAI/aspHI
    ecl136II
    bsp1286
    bsiHKAI
    bmyI                                                  tru9I                mspI
    banII                                                aflII/bfrI            hpaII
     aluI           bstXI      fokI  sfaNI mseI                               bsaWI
751 GGAGCTCATT TTCTTGCCAA AAGTTTGGAT GATGCCTTAA GACTTATTGA ACAACCGGAA TTGGCAAGTA
    CCTCGAGTAA AAGAACGGTT TTCAAACCTA CTACGGAATT CTGAATAACT TGTTGGCCTT AACCGTTCAT

                                                                               haeIII/palI
                                                                              haeI
                                                scrFI         scrFI
                                                mvaI                mvaI
                                                ecoRII        ecoRII
                                                dsaV   tfiI   dsaV
                                                bstNI    nlaIII     bstNI      ddeI
     accI nlaIII           mnlI                 apyI[dcm+]   hinfI  apyI[dcm+]
821 AAGTAGACAT GGTTTGGATA GTCGGAGGCA GTTCTGTTTA CCAGGAAGCC ATGAATCAAC CAGGCCACCT
    TTCATCTGTA CCAAACCTAT CAGCCTCCGT CAAGACAAAT GGTCCTTCGG TACTTAGTTG GTCCGGTGGA
```

FIG.32E

```
                           nlaIII
                       sau3AI
                       mboI/ndeII[dam-]
                       dpnI[dam+]                       maeII
  pleI                 dpnII[dam-]                    aflIII
  hinfI     maeIII alwI[dam-]  apoI          maeIII
891 TAGACTCTTT GTGACAAGGA TCATGCAGGA ATTTGAAAGT GACACGTTTT TCCCAGAAAT TGATTTGGGG
    ATCTGAGAAA CACTGTTCCT AGTACGTCCT TAAACTTTCA CTGTGCAAAA AGGGTCTTTA ACTAAACCCC

                              hgaI
                             hinlI/acyI
                             ahaII/bsaHI
                          scrFI
                          mvaI     mnlI
                          ecoRII
                          dsaV
                          bstNI    ecoNI
                          apyI[dcm+]      mnlI
              mnlI       bsaJI     bslI ddeI
961 AAATATAAAC CTCTCCCAGA ATACCCAGGC GTCCTCTCTG
    TTTATATTTG GAGAGGGTCT TATGGGTCCG CAGGAGAGAC
```

FIG.32F

```
        scrFI
        mvaI
        ecoRII
        dsaV
        bstNI
        apyI[dcm+]
     sau96I
     avaII                                                                            sfaNI
     asuI   mnlI        sfaNI                   accI      mboII                          mboII
1001 AGGTCCAGGA GGAAAAAGGC ATCAAGTATA AGTTTGAAGT CTACGAGAAG AAAGACTAAC AGGAAGATGC
     TCCAGGTCCT CCTTTTTCCG TAGTTCATAT TCAAACTTCA GATGCTCTTC TTTCTGATTG TCCTTCTACG
                                                                      ^END DHFR

                                                            nlaIII
                                                           styI
                                                           ncoI
                                               ppu10I      dsaI
                        mnlI      aluI nsiI/avaIII      bsaJI
1071 TTTCAAGTTC TCTGCTCCCC TCCTAAAGCT ATGCATTTTT ATAAGACCAT GGGACTTTTG
     AAAGTTCAAG AGACGAGGGG AGGATTTCGA TACGTAAAAA TATTCTGGTA CCCTGAAAAC
```

FIG.32G

```
                styI
                bsaJI
          sau3AI                fnu4HI
          mboI/ndeII[dam-]      aciI
          dpnI[dam+]            thaI
          dpnII[dam-]           fnuDII/mvnI tru9I
          alwI[dam-]            bstUI       mseI
          bstYI/xhoII           bsh1236I    aseI/asnI/vspI
1131 CTGGCTTTAG ATCCCCTTGG CTTCGTTAGA ACGCGGCTAC AATTAATACA TAACCTTATG TATCATACAC
     GACCGAAATC TAGGGGAACC GAAGCAATCT TGCGCCGATG TTAATTATGT ATTGGAATAC ATAGTATGTG

                                                                  sau96I
                                                                  avaII
                                                                  asuI
                                                            scrFI
                                                            mvaI
                                                            ecoRII
                                                            dsaV
                                                            bstNI
                                                            apyI[dcm+]
                                                      bslI  bsaJI
          maeIII                fokI
          hphI    scfI
1201 ATACGATTTA GGTGACACTA TAGATAACAT CCACTTTGCC TTTCTCTCCA CAGGTGTCCA CTCCCAGGTC
     TATGCTAAAT CCACTGTGAT ATCTATTGTA GGTGAAACGG AAAGAGAGGT GTCCACAGGT GAGGGTCCAG
```

FIG.32H

```
                                                        scrFI
                                                        nciI
                                                        mspI
                                                        hpaII
                                                        dsaV
                                                       xmaI/pspAI
                                                       smaI
                                                       scrFI
                                                       nciI
                                                       dsaV
                                                       cauII
                                                       bsaJI
                                                       avaI
                                                   sau3AI
                                                   mboI/ndeII[dam-]
                                                   dpnI[dam+]
                                                   dpnII[dam-]
                                       pleI       nlaIV cauII
                                       hinfI      bstYI/xhoII
                        scfI         taqI   rmaI  bamHI bsaJI
                        aluI  pstI   salI   maeI  alwI[dam-]
            mnlI       hindIII  bspMI hincII/hindII alwI[dam-]
            bsaJI    ddeI     bsgI    accI  xbaI mnlI  bsaJI
1271 CAACTGCACC TCGGTTCTAA GCTTCTGCAG GTCGACTCTA GAGGATCCCC
     GTTGACGTGG AGCCAAGATT CGAAGACGTC CAGCTGAGAT CTCCTAGGGG
```

FIG.32I

```
                    sau96I
                 aciI   haeIII/palI
               fnu4HI   asuI
                 bglI  nlaIII
                  sfiI  styI
                 eaeI   ncoI
                  cfrI   dsaI
          ecoRI   taqI haeIII/palI                    aluI
            apoI   claI/bsp106 bsaJI                fnu4HI
1321 GGGGAATTCA ATCGATGGCC GCCATGGCCC AACTTGTTTA TTGCAGCTTA      bbvI      maeIII
     CCCCTTAAGT TAGCTACCGG CGGTACCGGG TTGAACAAAT AACGTCGAAT ATTACCAATG TTTATTTCGT
                                      ^sv40 early poly A
                                                    rmaI
                                                 bsmI maeI
          sfaNI   apoI
1391 ATAGCATCAC AAATTTCACA AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGGTTTGT CCAAACTCAT
     TATCGTAGTG TTTAAAGTGT TTATTTCGTA AAAAAAGTGA CGTAAGATCA ACACCAAACA GGTTTGAGTA
```

FIG.32J

```
                                  sau3AI
                                  mboI/ndeII[dam-]
                                  dpnI[dam+]
                                  dpnII[dam-]
                                 pvuI/bspCI
                                 mcrI
                                taqI[dam-]  tru9I
                               claI/bsp106[dam-]
                              sau3AI       mseI
                              mboI/ndeII[dam-]
                              dpnI[dam+] xmnI
                              dpnII[dam-]  aseI/asnI/vspI
                     nlaIII  alwI[dam-]  asp700
1461 CAATGTATCT TATCATGTCT GGATCGATCG GGAATTAATT
     GTTACATAGA ATAGTACAGA CCTAGCTAGC CCTTAATTAA
                              sv40 origin^

                                                            rsaI
                      haeIII/palI                           csp6I
                     haeI                                 nlaIV
                  styI                                    kpnI
         fnu4HI ncoI                                      hgiCI
         bbvI   dsaI                                      banI
       hinPI    bsaJI                                     asp718     mnlI
       hhaI/cfoI nlaIII         mnlI       mnlI           acc65I   ddeI   aciI
1501 CGGCGCAGCA CCATGGCCTG AAATAACCTC TGAAAGAGGA ACTTGGTTAG GTACCTTCTG AGGCGGAAAG
     GCCGCGTCGT GGTACCGGAC TTTATTGGAG ACTTTCTCCT TGAACCAATC CATGGAAGAC TCCGCCTTTC
```

FIG.32K

FIG.32L

```
                                                                  nlaIV
                                                                 scrFI
                                                                 mvaI
                                                                 ecoRII
                                                                 dsaV
          aluI                                                   bstNI
         pvuII                                                   apyI[dcm+]
         nspBII                                                 bsaJI
1571 AACCAGCTGT GGAATGTGTG TCAGTTAGGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC AGAAGTATGC
     TTGGTCGACA CCTTACACAC AGTCAATCCC ACACCTTTCA GGGGTCCGAG GGGTCGTCCG TCTTCATACG

                                                                    nlaIV
      sfaNI                  scrFI        scrFI
          ppu10I                   mvaI                 mvaI
    nsiI/avaIII              ecoRII       ecoRII
   nlaIII                    dsaV         dsaV
        sphI                       bstNI                bstNI
        nspI                       apyI[dcm+]           apyI[dcm+]
        nspHI                     sexAI                bsaJI
1641 AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT
     TTTCGTACGT AGAGTTAATC AGTCGTTGGT CCACACCTTT CAGGGGTCCG AGGGGTCGTC CGTCTTCATA

             sfaNI
           ppu10I
           nsiI/avaIII
          nlaIII
         sphI
         nspI                                                                 aciI
         nspHI                                    aciI          aciI  fokI
1711 GCAAAGCATG CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC GCCCCTAACT
     CGTTTCGTAC GTAGAGTTAA TCAGTCGTTG GTATCAGGGC GGGGATTGAG GCGGGTAGGG CGGGGATTGA
```

```
                                                 nlaIII
                                                styI
                                                ncoI
        bsrI                               bslI dsaI
     asiI        aciI                      aciI bsaJI                              mnlI
1781 CCGCCCAGTT CCGCCCATTC TCCGCCCCAT GGCTGACTAA TTTTTTTTAT TTATGCAGAG
     GGCGGGTCAA GGCGGGTAAG AGGCGGGGTA CCGACTGATT AAAAAAAATA AATACGTCTC

                                                                                    styI
           fnu4HI                                                                   bsaJI
           bglI                                                                     blnI
          sfiI                                                                      avrII
          haeIII/palI                                                             haeIII/palI
        mnlI    mnlI     ddeI                                                     stuI
   haeIII/palI  bsaJI mnlI    aluI                               mnlI             haeI
      bsaJI aciI     haeIII/palI                              mnlI               mnlI
1841 GCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT TTTTGGAGGC
     CGGCTCCGGC GGAGCCGGAG ACTCGATAAG GTCTTCATCA CTCCTCCGAA AAAACCTCCG
```

FIG.32M

```
                                          aciI            hinPI
                                        haeIII/palI       hhaI/cfoI
                                       mcrI              thaI
                                       eagI/xmaIII/eclXI fnuDII/mvnI
                                 taqI  eaeI              bstUI              bspMI
                                xhoI  notI              hinPI             scfI
                                paeR7I cfrI   tru9I     hhaI/cfoI tru9I   pstI
                                avaI  fnu4HI  pacI     ascI      ahaIII/draI
     rmaI                      mnlI   aciI    mseI tru9I bsh1236I mseI    bsgI
     maeI              aluI maeIII  bsrBI fnu4HI    mseI bssHII  swaI     sse8387I
1901 CTAGGCTTTT GCAAAAAGCT GTTACCTCGA GCGGCCGCTT AATTAAGGCG CGCCATTTAA ATCCTGCAGG
     GATCCGAAAA CGTTTTTCGA CAATGGAGCT CGCCGGCGAA TTAATTCCGC GCGGTAAATT TAGGACGTCC
                     ^start pUC118
                             ^linearization linker inserted into Hpal site

                                                          scrFI
                                                          mvaI
                                                          ecoRII
                     haeIII/palI                          dsaV
                    eaeI                                  bstNI
   maeIII           cfrI                   maeIII         apyI[dcm+]         tru9I
         aluI     bsrI                maeII   bsrI       bsaJI   maeIII      mseI
1971 TAACAGCTTG GCACTGGCCG TCGTTTTACA ACGTCGTGAC TGGGAAAACC CTGGCGTTAC CCAACTTAAT
     ATTGTCGAAC CGTGACCGGC AGCAAAATGT TGCAGCACTG ACCCTTTTGG GACCGCAATG GGTTGAATTA
```

FIG.32N

```
                                                    sau3AI
                                                sau96I     mboI/ndeII[dam-]
                                            haeIII/pa1I
                                        asuI       dpnI[dam+]
                                            mnlI       dpnII[dam-]
                                    aluI        mboII    aciI  pvuI/bspCI
                                pvuII           earI/ksp632I   mcrI
                            fnu4HI      nspBII
                        bbvI    fokI                                hinPI
2041 CGCCTTGCAG CACATCCCCC CTTCGCCAGC TGGCGTAATA GCGAAGAGGC CCGCACCGAT  hhaI/cfoI
     GCGGAACGTC GTGTAGGGGG GAAGCGGTCG ACCGCATTAT CGCTTCTCCG GGCGTGGCTA nlaIV
                                                                        narI
                                                                        kasI
                                                                        hinlI/acyI
                                                                        hgiCI
                                                                        haeII     aciI
                                                                        banI    sfaNI
                                                                    ahaII/bsaHI
                                                    bglI                    sfaNI
2101 CGCCCTTCCC AACAGTTGCG TAGCCTGAAT GGCGAATGGC GCCTGATGCG GTATTTTCTC CTTACGCATC
     GCGGGAAGGG TTGTCAACGC ATCGGACTTA CCGCTTACCG CGGACTACGC CATAAAGAG GAATGCGTAG
```

FIG.32O

```
                                                        hinPI
                                                       thaI
                                                       fnuDII/mvnI
                                                       bstUI scfI     hinPI
                                                       bsh1236I       hhaI/cfoI
                                                   rsaI hhaI/cfoI  fnu4HI
     aciI          aciI   maeII                    csp6I   bslI    aciI
2171 TGTGCGGTAT TTCACACCGC ATACGTCAAA GCAACCATAG TACGCGCCCT GTAGCGGCGC
     ACACGCCATA AAGTGTGGCG TATGCAGTTT CGTTGGTATC ATGCGCGGGA CATCGCCGCG

         fnu4HI                   fnu4HI
        thaI                    hinPI
        fnuDII/mvnI             hhaI/cfoI                          hinPI
        bctUI                 thaI                                 hhaI/cfoI
       hinPI aciI             fnuDII/mvnI                      rmaI
       hhaI/cfoI              bstUI                        hinPI   haeII
   tru9I aciI                 bsh1236I      aciI           hhaI/cfoI   bsrBI
   mseI bsh1236I           maeIII bbvI  maeIII            haeII maeI   aciI
2231 ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT
     TAATTCGCGC CGCCCACACC ACCAATGCGC GTCGCACTGG CGATGTGAAC GGTCGCGGGA TCGCGGGCGA

                                                                    nlaIV
                                            mspI                   hgiJII
                                            hpaII                  bsp1286
                                           naeI                    bmyI
                mboII                 maeII cfr10I          aluI    banII
2301 CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG GCTTTCCCCG TCAAGCTCTA AATCGGGGGC
     GGAAAGCGAA AGAAGGGAAG GAAAGAGCGG TGCAAGCGGC CGAAAGGGGC AGTTCGAGAT TTAGCCCCCG
```

FIG.32P

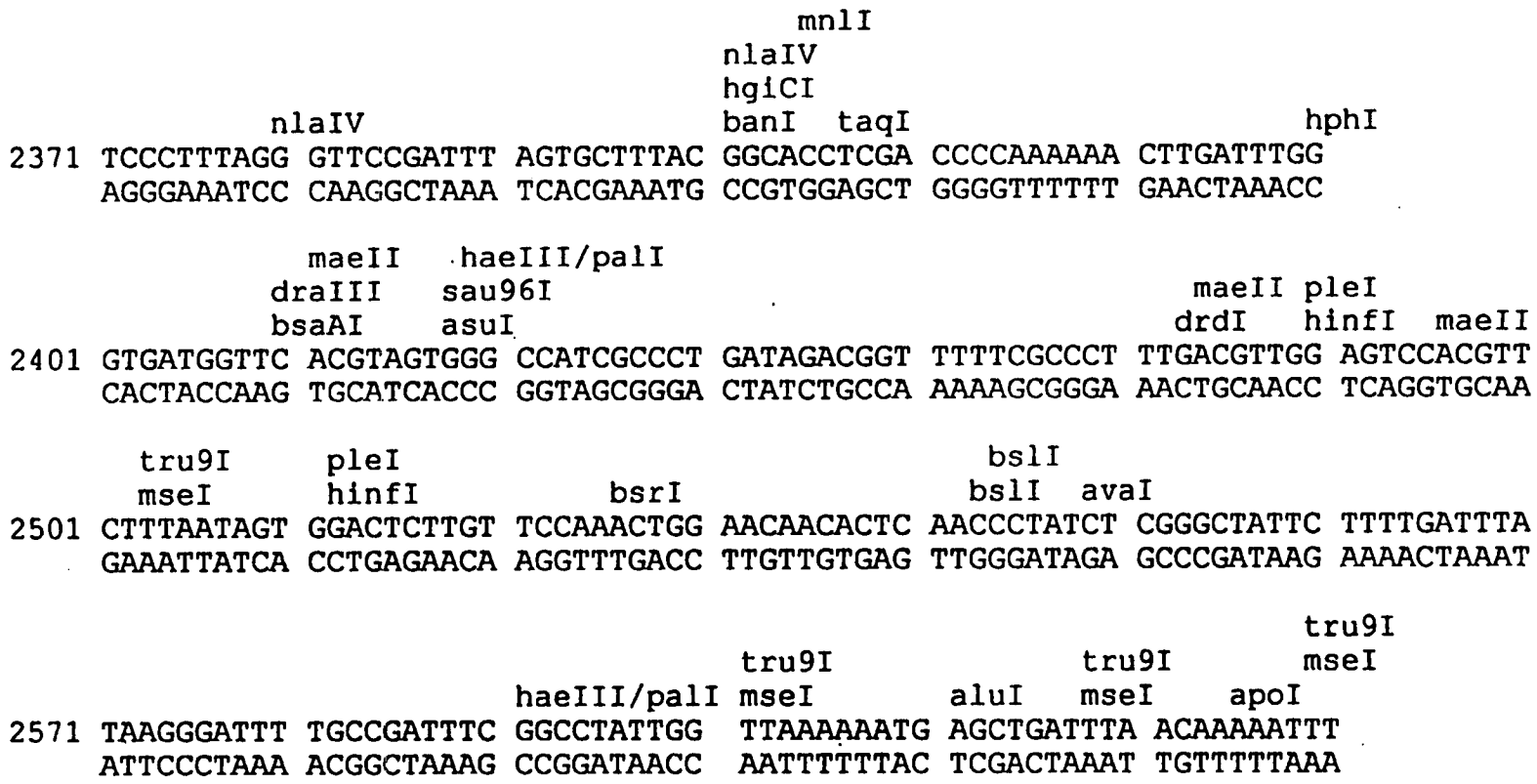

mnlI
nlaIV
hgiCI
nlaIV
banI taqI
hphI
2371 TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCAAAAAA CTTGATTTGG
AGGGAAATCC CAAGGCTAAA TCACGAAATG CCGTGGAGCT GGGGTTTTTT GAACTAAACC
maeII haeIII/palI
draIII sau96I
maeII pleI
bsaAI asuI
drdI hinfI maeII
2401 GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT TTGACGTTGG AGTCCACGTT
CACTACCAAG TGCATCACCC GGTAGCGGGA CTATCTGCCA AAAAGCGGGA AACTGCAACC TCAGGTGCAA
tru9I pleI
mseI hinfI
bslI
bsrI
bslI avaI
2501 CTTTAATAGT GGACTCTTGT TCCAAACTGG AACAACACTC AACCCTATCT CGGGCTATTC TTTTGATTTA
GAAATTATCA CCTGAGAACA AGGTTTGACC TTGTTGTGAG TTGGGATAGA GCCCGATAAG AAAACTAAAT
tru9I
tru9I
tru9I
mseI
haeIII/palI
mseI
aluI
mseI
apoI
2571 TAAGGGATTT TGCCGATTTC GGCCTATTGG TTAAAAAATG AGCTGATTTA ACAAAAATTT
ATTCCCTAAA ACGGCTAAAG CCGGATAACC AATTTTTTAC TCGACTAAAT TGTTTTTAAA

FIG.32Q

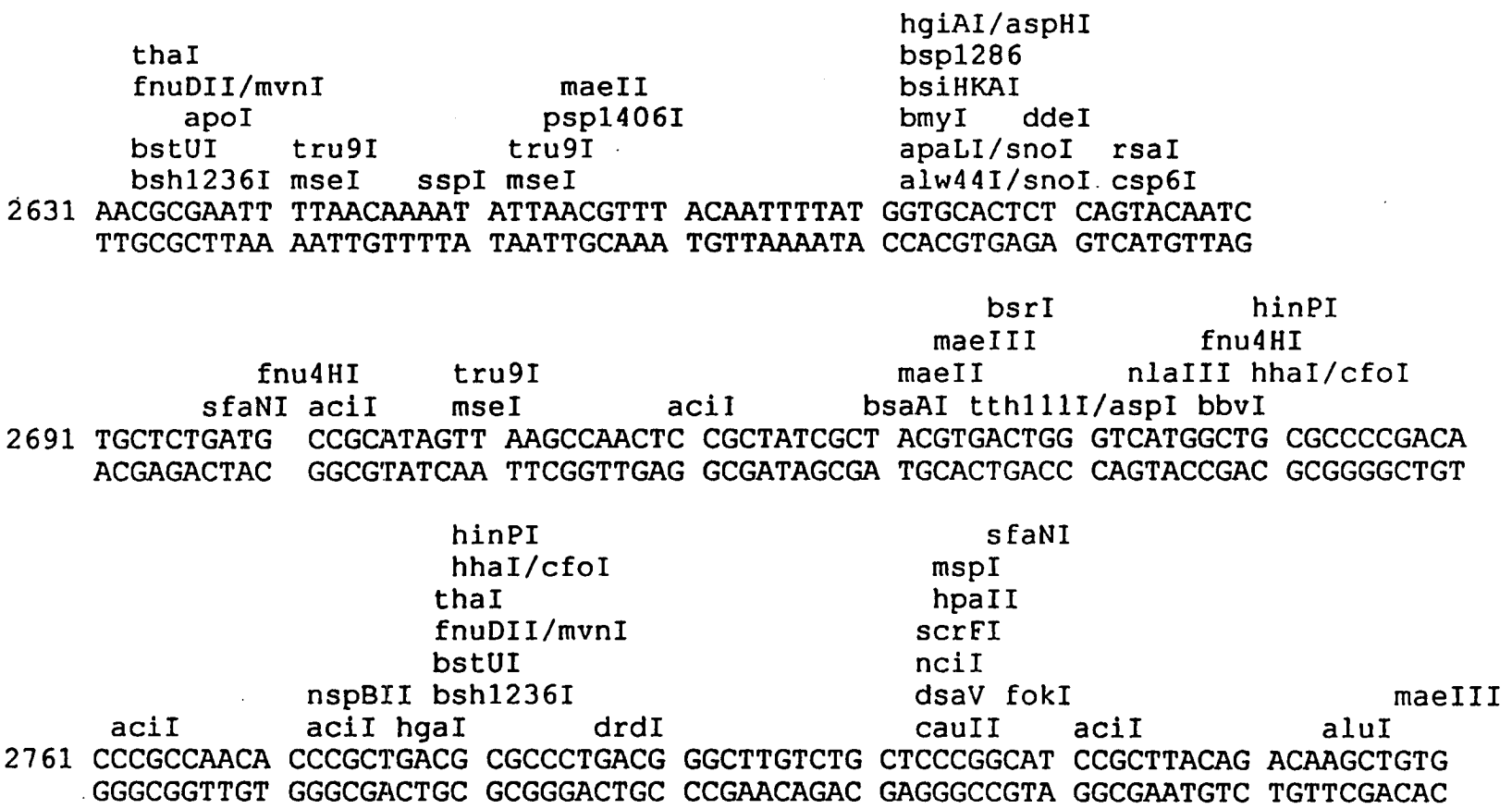

hgiAI/aspHI
thaI
bsp1286
fnuDII/mvnI
maeII
bsiHKAI
apoI
psp1406I
bmyI ddeI
bstUI tru9I tru9I
apaLI/snoI rsaI
bsh1236I mseI sspI mseI
alw44I/snoI csp6I
2631 AACGCGAATT TTAACAAAAT ATTAACGTTT ACAATTTTAT GGTGCACTCT CAGTACAATC
TTGCGCTTAA AATTGTTTTA TAATTGCAAA TGTTAAAATA CCACGTGAGA GTCATGTTAG
bsrI hinPI
maeIII fnu4HI
fnu4HI tru9I maeII nlaIII hhaI/cfoI
sfaNI aciI mseI aciI bsaAI tth111I/aspI bbvI
2691 TGCTCTGATG CCGCATAGTT AAGCCAACTC CGCTATCGCT ACGTGACTGG GTCATGGCTG CGCCCCGACA
ACGAGACTAC GGCGTATCAA TTCGGTTGAG GCGATAGCGA TGCACTGACC CAGTACCGAC GCGGGGCTGT
hinPI sfaNI
hhaI/cfoI mspI
thaI hpaII
fnuDII/mvnI scrFI
bstUI nciI
nspBII bsh1236I dsaV fokI maeIII
aciI aciI hgaI drdI cauII aciI aluI
2761 CCCGCCAACA CCCGCTGACG CGCCCTGACG GGCTTGTCTG CTCCCGGCAT CCGCTTACAG ACAAGCTGTG
GGGCGGTTGT GGGCGACTGC GCGGGACTGC CCGAACAGAC GAGGGCCGTA GGCGAATGTC TGTTCGACAC

FIG.32R

```
                                                                    thaI
                                                                    fnuDII/mvnI
          scrFI                                                     bstUI
          nciI                                                      bsh1236I
          mspI                                                     hinPI
          hpaII      nspI                                          hhaI/cfoI
          dsaV       nspHI                                       thaI mnlI
       esp3I        fnu4HI                                       fnuDII/mvnI
        bsmAI       bbvI                                         bstUI
      bslI   cauII aluI nlaIII     mnlI     hphI       hphI      bsh1236I
2831 ACCGTCTCCG GGAGCTGCAT GTGTCAGAGG TTTTCACCGT CATCACCGAA ACGCGCGAGG CAGTATTCTT
     TGGCAGAGGC CCTCGACGTA CACAGTCTCC AAAAGTGGCA GTAGTGGCTT TGCGCGCTCC GTCATAAGAA

                 mnlI
                haeIII/palI
     mboII       sau96I                                        nlaIII
     bpuAI       asuI                                  tru9I  rcaI
     bbsI     eco0109I/draII                           mseI   bspHI
2901 GAAGACGAAA GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG
     CTTCTGCTTT CCCGGAGCAC TATGCGGATA AAAATATCCA ATTACAGTAC

                                                                  aciI
                                                                 thaI
                                                                 fnuDII/mvnI
                     hinII/acyI                                  bstUI
                     ahaII/bsaHI                                 bsh1236I
                     aatII                                      hinPI
                   ddeI maeII                                   hhaI/cfoI
2951 ATAATAATGG TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG GAAATGTGCG
     TATTATTACC AAAGAATCTG CAGTCCACCG TGAAAAGCCC CTTTACACGC
```

FIG.32S

```
                                                                  bsmAI
                                                               rcaI
                                                         bsrBI nlaIII
                                                            aciI bspHI
      nlaIV
3001  CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA
      GCCTTGGGGA TAAACAAATA AAAAGATTTA TGTAAGTTTA TACATAGGCG AGTACTCTGT TATTGGGACT

                                  mboII
                        sspI      earI/ksp632I
3071  TAAATGCTTC AATAATATTG AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTCGCC CTTATTCCCT
      ATTTACGAAG TTATTATAAC TTTTTCCTTC TCATACTCAT AAGTTGTAAA GGCACAGCGG GAATAAGGGA

            fnu4HI
              aciI                    hphI                hphI          sfaNI
3141  TTTTTGCGGC ATTTTGCCTT CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG
      AAAAACGCCG TAAAACGGAA GGACAAAAAC GAGTGGGTCT TTGCGACCAC TTTCATTTTC

                        hgiAI/aspHI
                           bsp1286                        sau3AI
                  sau3AI   bsiHKAI                        mboI/ndeII[dam-]
                  mboI/ndeII[dam-]                        dpnI[dam+]
                  dpnI[dam+] bmyI                         dpnII[dam-]
                  dpnII[dam-]                          bstYI/xhoII
            mboII[dam-]    apaLI/snoI        taqI       alwI[dam-]   aciI
              eco57I       alw44I/snoI       maeIII     bsrI         nspBII
3201  ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA
      TACGACTTCT AGTCAACCCA CGTGCTCACC CAATGTAGCT TGACCTAGAG TTGTCGCCAT
```

FIG.32T

```
         sau3AI
         mboI/ndeII[dam-]             maeII
         dpnI[dam+]                  psp1406I          hgiAI/aspHI
         dpnII[dam-]                xmnI               bsp1286  tru9I
         alwI[dam-]                 asp700             bsiHKAI  mseI
        bstYI/xhoII              mboII                 bmyI    ahaIII/draI
3261    AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC
        TCTAGGAACT CTCAAAAGCG GGGCTTCTTG CAAAAGGTTA CTACTCGTGA AAATTTCAAG

                                           scrFI
                   aciI                    nciI
                  thaI                     mspI
                  fnuDII/mvnI              hpaII
                  bstUI                    dsaV
                  bsh1236I             hinlI/acyI
                hinPI                  hgaI cauII                  aciI
                hhaI/cfoI              ahaII/bsaHI    bcgI  mcrI  fnu4HI
3321    TGCTATGTGG CGCGGTATTA TCCCGTGATG ACGCCGGGCA AGAGCAACTC GGTCGCCGCA
        ACGATACACC GCGCCATAAT AGGGCACTAC TGCGGCCCGT TCTCGTTGAG CCAGCGGCGT

                                   rsaI
                                   csp6I   bsrI
                ddeI              scaI  hphI  maeIII     sfaNI     fokI
3381    TACACTATTC TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG
        ATGTGATAAG AGTCTTACTG AACCAACTCA TGAGTGGTCA GTGTCTTTTC GTAGAATGCC
```

FIG.32U

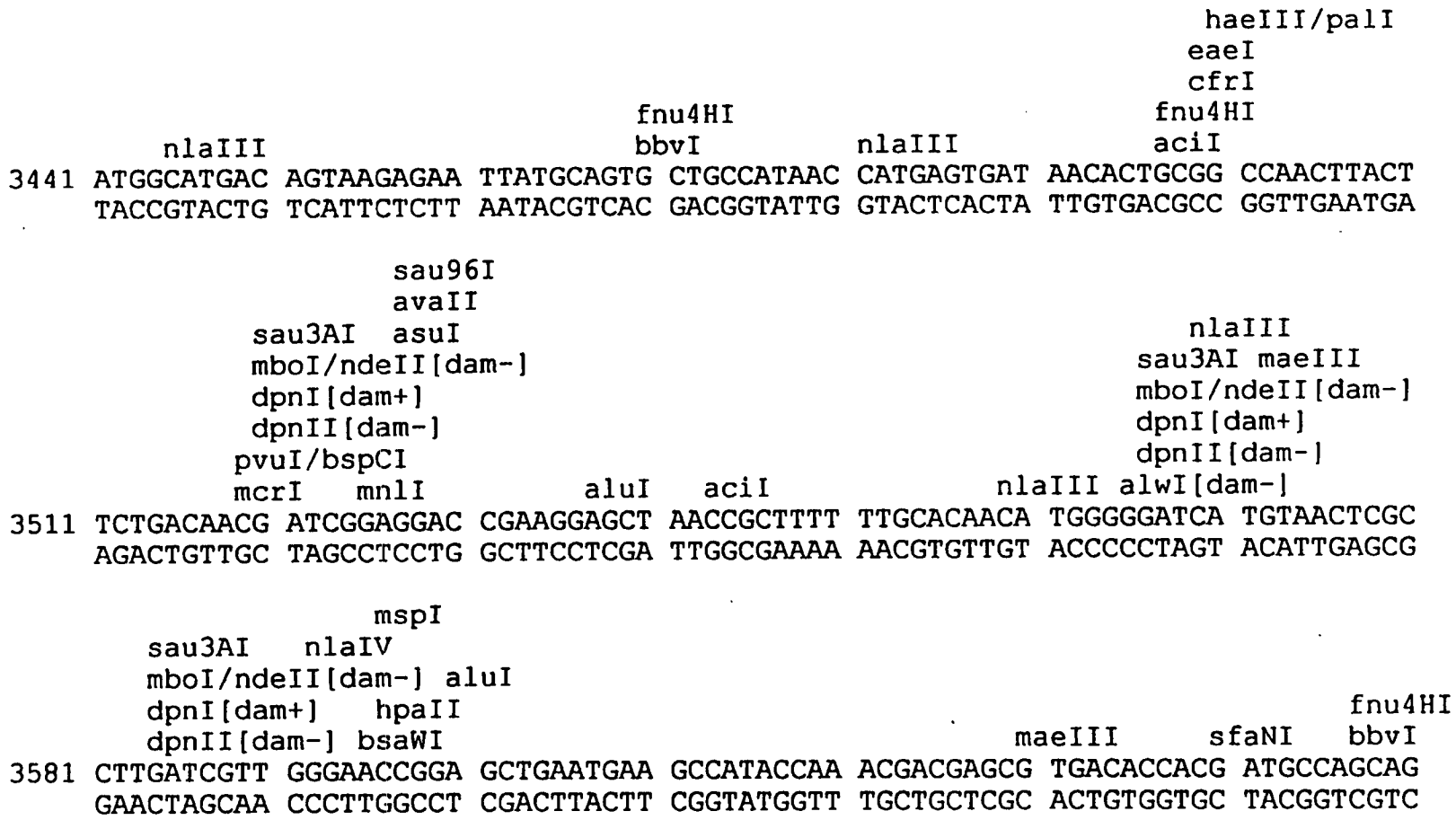

haeIII/palI
eaeI
cfrI
fnu4HI
fnu4HI
nlaIII
bbvI
nlaIII
aciI
3441 ATGGCATGAC AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT
TACCGTACTG TCATTCTCTT AATACGTCAC GACGGTATTG GTACTCACTA TTGTGACGCC GGTTGAATGA
sau96I
avaII
sau3AI asuI
mboI/ndeII[dam-]
dpnI[dam+]
dpnII[dam-]
pvuI/bspCI
mcrI mnlI aluI aciI
nlaIII
sau3AI maeIII
mboI/ndeII[dam-]
dpnI[dam+]
dpnII[dam-]
nlaIII alwI[dam-]
3511 TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC
AGACTGTTGC TAGCCTCCTG GCTTCCTCGA TTGGCGAAAA AACGTGTTGT ACCCCCTAGT ACATTGAGCG
mspI
sau3AI nlaIV
mboI/ndeII[dam-] aluI
dpnI[dam+] hpaII
dpnII[dam-] bsaWI
fnu4HI
maeIII sfaNI bbvI
3581 CTTGATCGTT GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCCAGCAG
GAACTAGCAA CCCTTGGCCT CGACTTACTT CGGTATGGTT TGCTGCTCGC ACTGTGGTGC TACGGTCGTC

FIG.32V

```
                              hinPI                                                  mspI
                              hhaI/cfoI                                              hpaII
                             mstI                                                   scrFI
                             aviII/fspI     bsrI                              aluI  nciI
                         maeII            tru9I                              rmaI    dsaV
                        psp1406I          mseI                               maeI    cauII
3651 CAATGGCAAC AACGTTGCGC AAACTATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC
     GTTACCGTTG TTGCAACGCG TTTGATAATT GACCGCTTGA TGAATGAGAT CGAAGGGCCG

                                                                      bglI
                                                                     sau96I
      tru9I       fokI                              sau96I           haeIII/palI
      mseI     bsrI        aciI                     avaII      hinPI asuI    mspI
     aseI/asnI/vspI     mnlI                        asuI       hhaI/cfoI     hpaII
3711 AACAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC TTCCGGCTGG
     TTGTTAATTA TCTGACCTAC CTCCGCCTAT TTCAACGTCC TGGTGAAGAC GCGAGCCGGG AAGGCCGACC

                              mspI                   thaI
                              hpaII                  fnuDII/mvnI
                             cfr10I                  bstUI
                          nlaIV hphI             bsmAI aciI         fnu4HI
                         gsuI/bpmI              bsaI  bsh1236I      bbvI
3781 CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA TCATTGCAGC
     GACCAAATAA CGACTATTTA GACCTCGGCC ACTCGCACCC AGAGCGCCAT AGTAACGTCG
```

FIG.32W

```
      sau96I                                                    pleI
        asuI                                                    hinfI
       nlaIV                                        eam1105I
    bsrI haeIII/palI      mnlI
3841 ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC
     TGACCCCGGT CTACCATTCG GGAGGGCATA GCATCAATAG ATGTGCTGCC CCTCAGTCCG

                                ddeI
                             sau3AI        nlaIV
                          mboI/ndeII[dam-]
                          dpnI[dam+]    hgiCI          tru9I
                       fokI dpnII[dam-]  banI mnlI     mseI
3901 AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG
     TTGATACCTA CTTGCTTTAT CTGTCTAGCG ACTCTATCCA CGGAGTGACT AATTCGTAAC

                                                 tru9I
                                                 mseI          tru9I
                                              ahaIII/draI      mseI
   maeIII
3961 GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC TTCATTTTTA
     CATTGACAGT CTGGTTCAAA TGAGTATATA TGAAATCTAA CTAAATTTTG AAGTAAAAAT

          rmaI      sau3AI
       sau3AI hphI  mboI/ndeII[dam-]
       mboI/ndeII[dam-]
       dpnI[dam+]   dpnI[dam+]
       dpnII[dam-]  dpnII[dam-]                          maeII
    tru9I bstYI/xhoII  alwI[dam-]     nlaIII           tru9I
     mseI  alwI[dam-]  bstYI/xhoII    rcaI             mseI
   ahaIII/draI maeI  mboII[dam-]      bspHI
4021 ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTCG
     TAAATTTTCC TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTTT AGGGAATTGC ACTCAAAAGC
```

FIG.32X

FIG.32Y

```
                                                    sau3AI
                                                    mboI/ndeII[dam-]
                                                    dpnI[dam+]   sau3AI
                                                    dpnII[dam-] mboI/ndeII[dam-]
                                                   bstYI/xhoII  dpnI[dam+]
                                        sau3AI  alwI[dam-]    dpnII[dam-]
                                        mboI/ndeII[dam-]      alwI[dam-]
                 hgaI                   dpnI[dam+] mboII[dam-]
               ddeI                     dpnII[dam-]          bstYI/xhoII
4091 TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT
     AAGGTGACTC GCAGTCTGGG GCATCTTTTC TAGTTTCCTA GAAGAACTCT AGGAAAAAAA

        thaI
        fnuDII/mvnI
        bstUI
        bsh1236I
      hinPI       fnu4HI                                  aciI
      hhaI/cfoI   bbvI                          aciI     nspBII
4151 CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTTGTTTG
     GACGCGCATT AGACGACGAA CGTTTGTTTT TTTGGTGGCG ATGGTCGCCA CCAAACAAAC

        sau3AI
        mboI/ndeII[dam-]
        dpnI[dam+]
        dpnII[dam-]
      alwI[dam-]
    mspI                                        bsrI              hinPI
    hpaII       aluI                          maeIII   eco57I     hhaI/cfoI
4211 CCGGATCAAG AGCTACCAAC TCTTTTTCCG AAGGTAACTG GCTTCAGCAG AGCGCAGATA CCAAATACTG
     GGCCTAGTTC TCGATGGTTG AGAAAAAGGC TTCCATTGAC CGAAGTCGTC TCGCGTCTAT GGTTTATGAC
```

```
          rmaI                  haeIII/palI
          maeI      bslI      haeI                  scfI       aciI       mnlI
4281 TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT
     AGGAAGATCA CATCGGCATC AATCCGGTGG TGAAGTTCTT GAGACATCGT GGCGGATGTA TGGAGCGAGA

                                                           scrFI
                                                           nciI
                                 fnu4HI                    mspI
                         alwNI   bbvI                      hpaII
                        bsrI  fnu4HI                       dsaV      pleI
                   maeIII     bbvI     bsrI                cauII     hinfI
4351 GCTAATCCTG TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA
     CGATTAGGAC AATGGTCACC GACGACGGTC ACCGCTATTC AGCACAGAAT GGCCCAACCT GAGTTCTGCT

                                                     hgiAI/aspHI
                              nspBII                 bsp1286
            mspI             fnu4HI                  bsiHKAI
             hpaII           bbvI mcrI               bmyI
            bsaWI          hinPI aciI                apaLI/snoI
         maeIII            hhaI/cfoI                 alw44I/snoI   aluI
4421 TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA
     ATCAATGGCC TATTCCGCGT CGCCAGCCCG ACTTGCCCCC CAAGCACGTG TGTCGGGTCG AACCTCGCTT

                                                      hinPI
                                                      hhaI/cfoI
                      ddeI       scfI                haeII
4491 CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA
     GCTGGATGTG GCTTGACTCT ATGGATGTCG CACTCGTAAC TCTTTCGCGG TGCGAAGGGC TTCCCTCTTT
```

FIG.32Z-1

```
                                                                          scrFI
                                                                          mvaI
                                                                          ecoRII     mvaI
               mspI                                                       dsaV       ecoRII
               hpaII                                                      bstNI
               bslI       fnu4HI                         hinPI mnlI       bsaJI
     aciI      bsaWI      aciI                           hhaI/cfoI  aluI  apyI[dcm+]
4561 GGCGGACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC
     CCGCCTGTCC ATAGGCCATT CGCCGTCCCA GCCTTGTCCT CTCGCGTGCT CCCTCGAAGG TCCCCCTTTG

     scrFI

     dsaV
     bstNI                                                          taqI
     apyI[dcm+]                                mnlI  drdI   hgaI              sfaNI
4631 GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT
     CGGACCATAG AAATATCAGG ACAGCCCAAA GCGGTGGAGA CTGAACTCGC AGCTAAAAAC ACTACGAGCA

                                                 haeIII/palI
                                               fnu4HI
                                               aciI
                                              thaI bslI
                                              fnuDII/mvnI
            nlaIV                             bstUI
          aciI                                bsh1236I
4701 CAGGGGGGCG GAGCCTATGG AAAAACGCCA GCAACGCGGC
     GTCCCCCCGC CTCGGATACC TTTTTGCGGT CGTTGCGCCG
```

FIG.32Z-2

```
                    haeIII/palI
                 scrFI
                 mvaI bslI
                 ecoRII
                 dsaV                    nlaIII
                 bstNI                        nspI
                 apyI[dcm+]    haeIII/palI    nspHI                     tfiI
            nlaIV  haeI       haeI            aflIII                    hinfI
4741 CTTTTTACGG TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT
     GAAAAATGCC AAGGACCGGA AAACGACCGG AAAACGAGTG TACAAGAAAG GACGCAATAG GGGACTAAGA

                                                           fnu4HI
                                                           bbvI
                                                  bsrBI   aciI
                 aciI           aluI              aciI   fnu4HI       mcrI
4811 GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCGCAG CCGAACGACC
     CACCTATTGG CATAATGGCG GAAACTCACT CGACTATGGC GAGCGGCGTC GGCTTGCTGG

                                               hinPI
      fnu4HI                                  haeII
      bbvI pleI                           sapI hhaI/cfoI
     hinPI  hinfI                         mboII                     mnlI
     hhaI/cfoI               mnlI  aciI   earI/ksp632I             aciI
4871 GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC
     CTCGCGTCGC TCAGTCACTC GCTCCTTCGC CTTCTCGCGG GTTATGCGTT TGGCGGAGAG
```

FIG.32Z-3

```
      thaI
      fnuDII/mvnI
      bstUI
      bsh1236I
      hinPI
      hhaI/cfoI
      thaI
      fnuDII/mvnI                tru9I  aluI
      bstUI                                 pvuII
      bsh1236I haeIII/palI
      bslI     eaeI  tfiI aseI/asnI/vspI                    bsrI     aciI
      aciI     cfrI  hinfI  mseI  nspBII
4931  CCCGCGCGTT GGCCGATTCA TTAATCCAGC TGGCACGACA GGTTTCCCGA CTGGAAAGCG
      GGGCGCGCAA CCGGCTAAGT AATTAGGTCG ACCGTGCTGT CCAAAGGGCT GACCTTTCGC

                                                      scrFI
                                                      mvaI
                                                      ecoRII
                                                      dsaV
                                                nlaIV  bstNI
                                 tru9I          hgiCI  apyI[dcm+]
                hinPI            mseI    maeIII       banI bsaJI
                hhaI/cfoI aseI/asnI/vspI  mnlI
4991  GGCAGTGAGC GCAACGCAAT TAATGTGAGT TACCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC
      CCGTCACTCG CGTTGCGTTA ATTACACTCA ATGGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAAATACG
```

FIG.32Z-4

```
        mspI                                     aciI
         hpaII                                bsrBI                                           aluI       nlaIII
5061 TTCCGGCTCG TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTCACACA GGAAACAGCT ATGACCATGA
     AAGGCCGAGC ATACAACACA CCTTAACACT CGCCTATTGT TAAAGTGTGT CCTTTGTCGA TACTGGTACT

             tru9I
             mseI
            aseI/asnI/vspI
          xmnI
          asp700
5131 TTACGAATTA A
     AATGCTTAAT T
```

FIG.32Z-5

>length: 5141

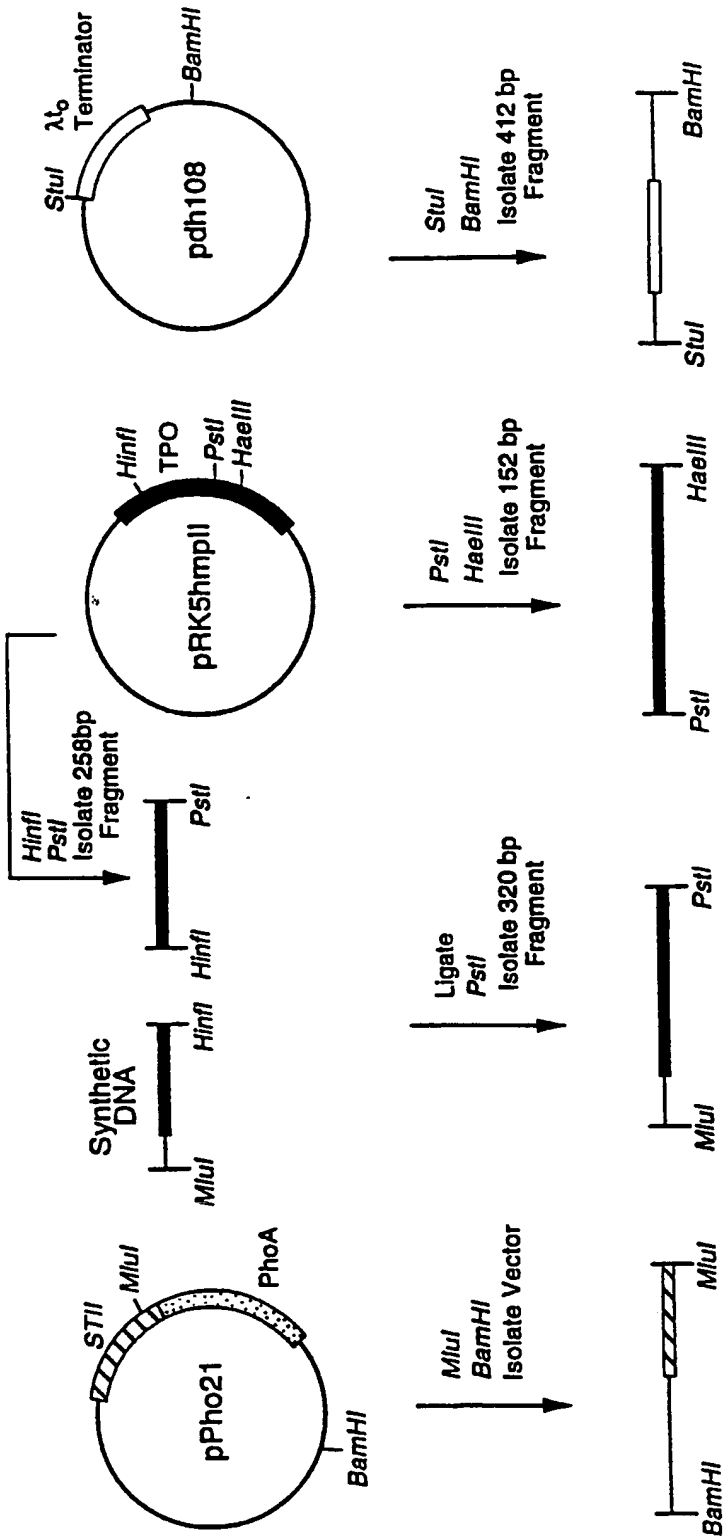

pPho21
STII
MluI
PhoA
BamHI
MluI
BamHI
Isolate Vector
MluI
BamHI
Synthetic DNA
HinfI
MluI
Ligate
PstI
Isolate 320 bp Fragment
PstI
MluI
HinfI
PstI
Isolate 258bp Fragment
HinfI
PstI
pRK5hmpII
HinfI
TPO
PstI
HaeIII
PstI
HaeIII
Isolate 152 bp Fragment
HaeIII
PstI
pdh108
StuI
λt0 Terminator
BamHI
StuI
BamHI
Isolate 412 bp Fragment
BamHI
StuI

FIG.33A

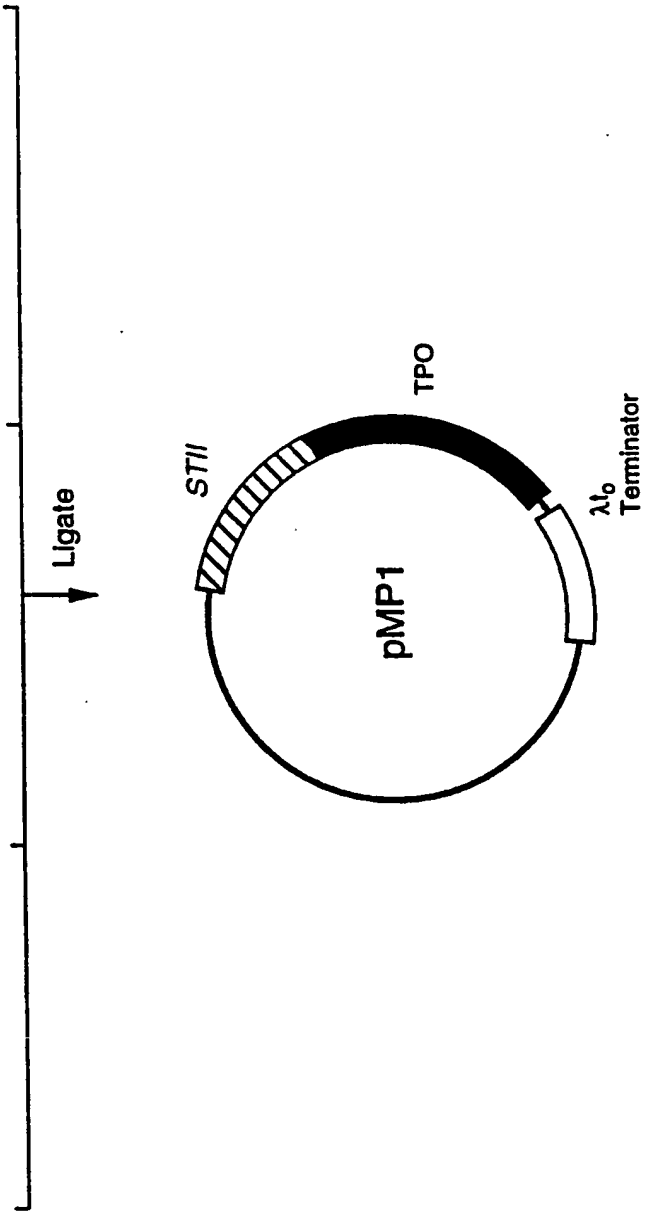


FIG.33B

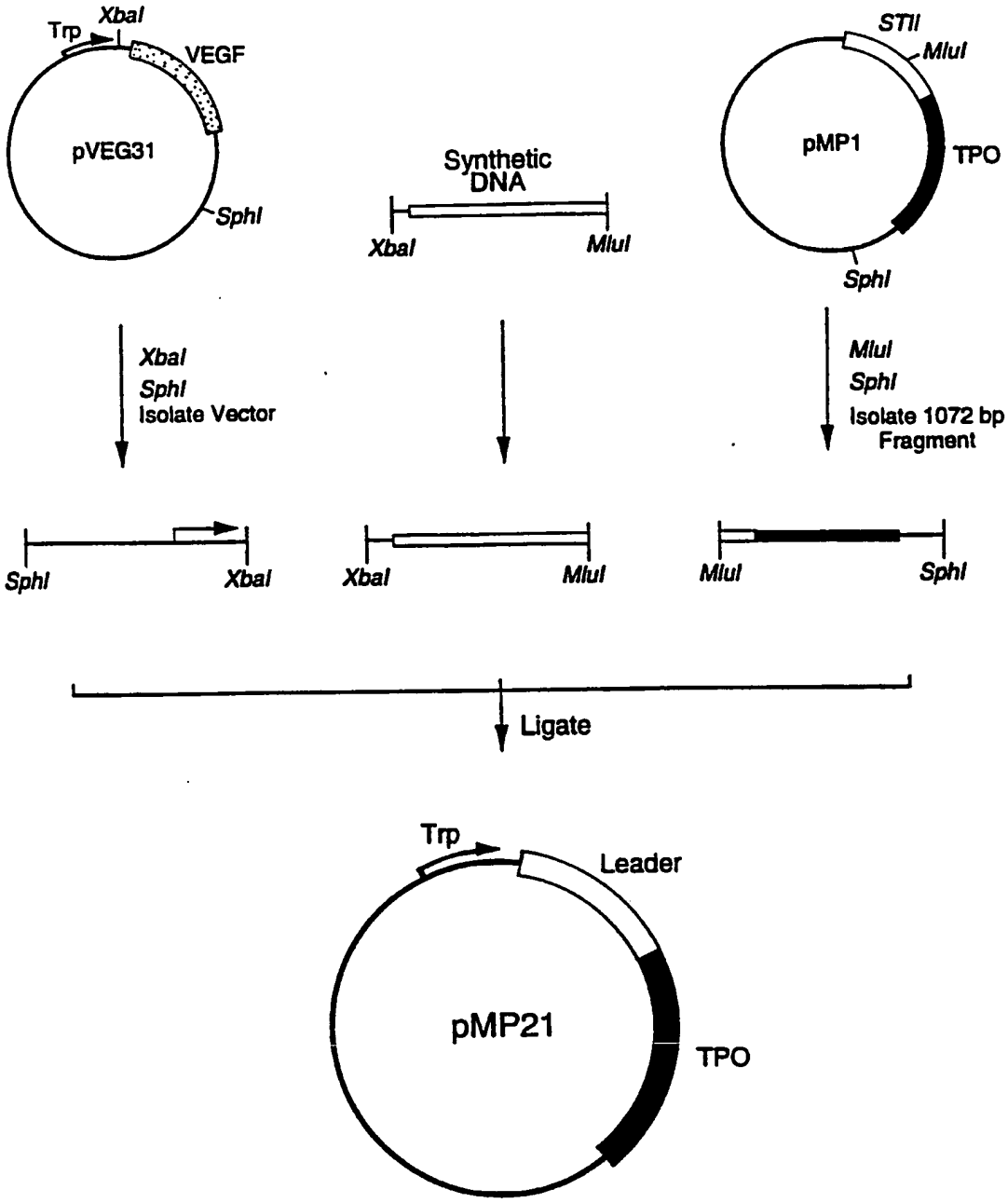

Trp
XbaI
VEGF
pVEG31
SphI
Synthetic DNA
XbaI
MluI
STII
MluI
pMP1
TPO
SphI
XbaI
SphI
Isolate Vector
MluI
SphI
Isolate 1072 bp Fragment
SphI
XbaI
XbaI
MluI
MluI
SphI
Ligate
Trp
Leader
pMP21
TPO

FIG.34

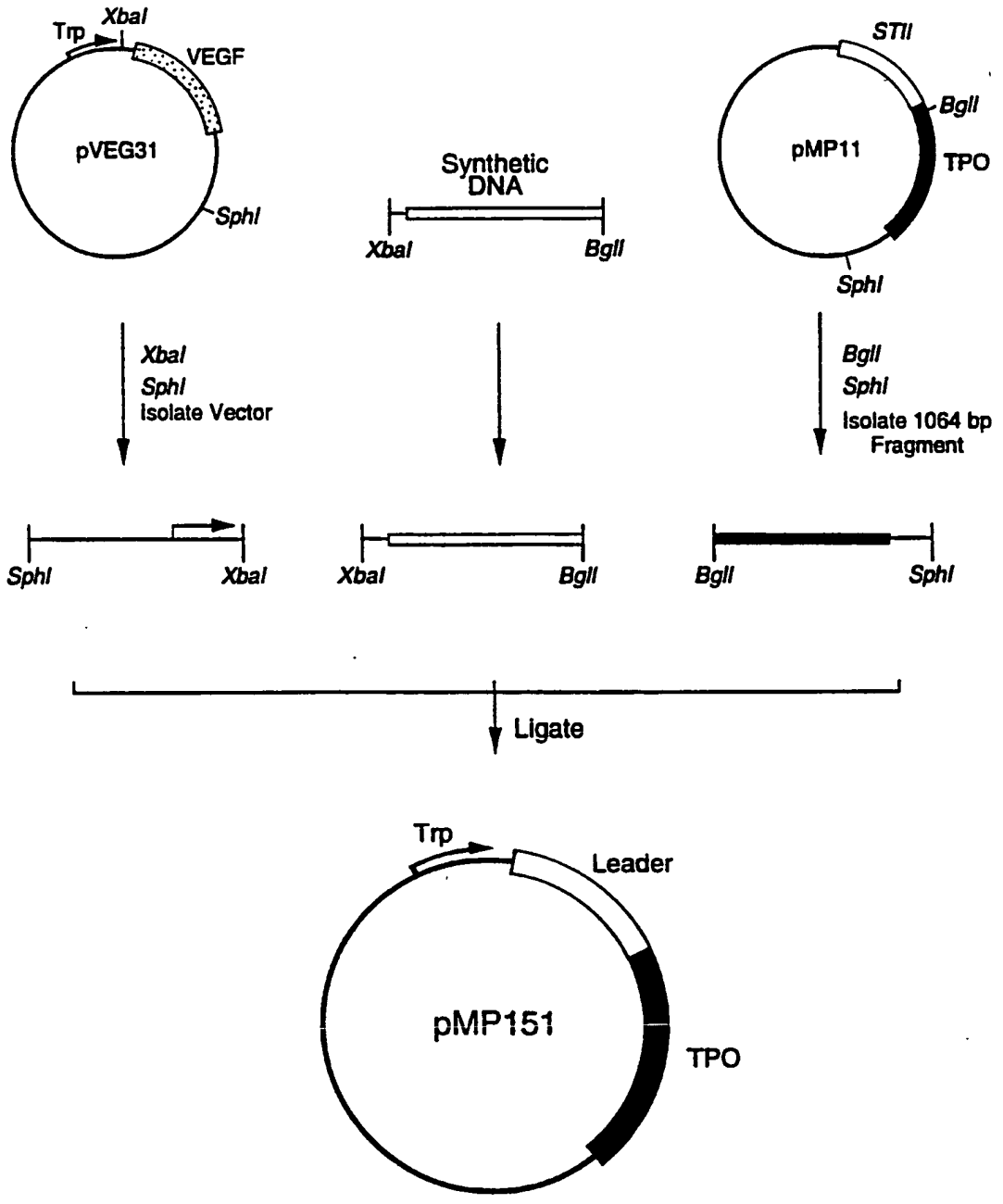

Trp
XbaI
VEGF
pVEG31
SphI
Synthetic DNA
XbaI
BglI
STII
BglI
pMP11
TPO
SphI
XbaI
SphI
Isolate Vector
BglI
SphI
Isolate 1064 bp Fragment
SphI
XbaI
XbaI
BglI
BglI
SphI
Ligate
Trp
Leader
pMP151
TPO

FIG.35

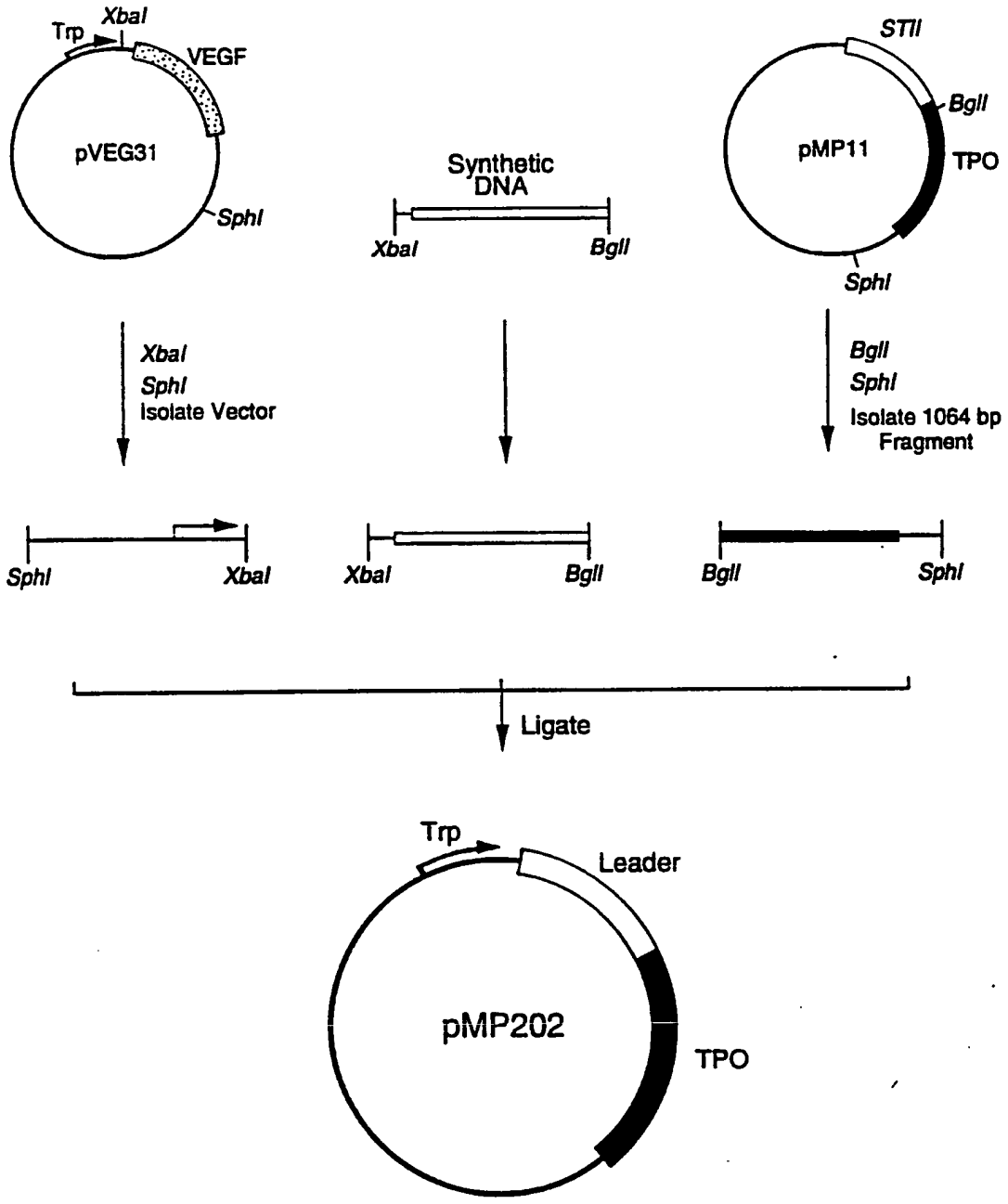

Trp
XbaI
VEGF
pVEG31
SphI
Synthetic DNA
XbaI
BglI
STII
BglI
pMP11
TPO
SphI
XbaI
SphI
Isolate Vector
BglI
SphI
Isolate 1064 bp Fragment
SphI
XbaI
XbaI
BglI
BglI
SphI
Ligate
Trp
Leader
pMP202
TPO

FIG.36

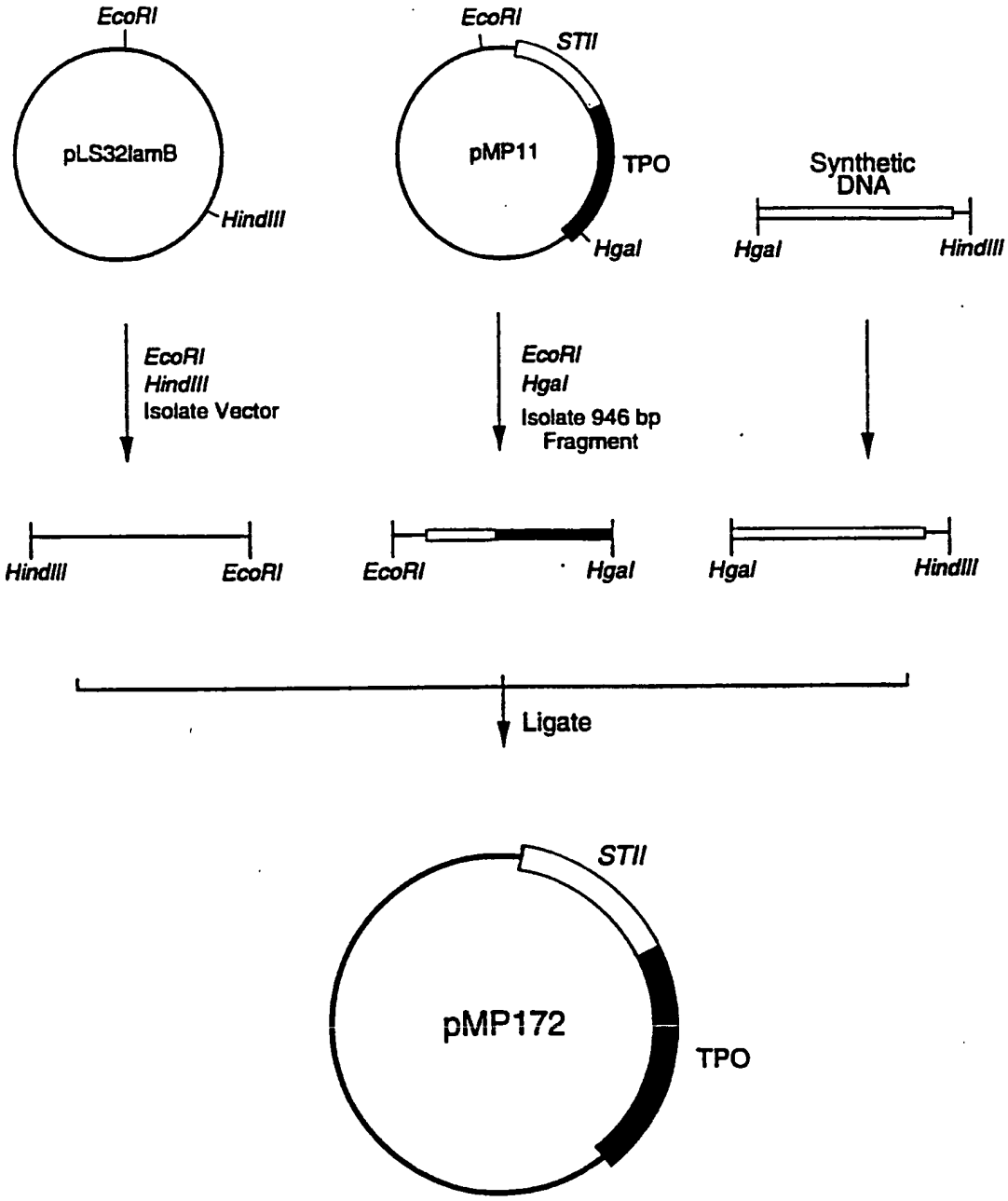

EcoRI
pLS32lamB
HindIII
EcoRI
STII
pMP11
TPO
HgaI
Synthetic DNA
HgaI
HindIII
EcoRI
HindIII
Isolate Vector
EcoRI
HgaI
Isolate 946 bp Fragment
HindIII
EcoRI
EcoRI
HgaI
HgaI
HindIII
Ligate
STII
pMP172
TPO

FIG.37

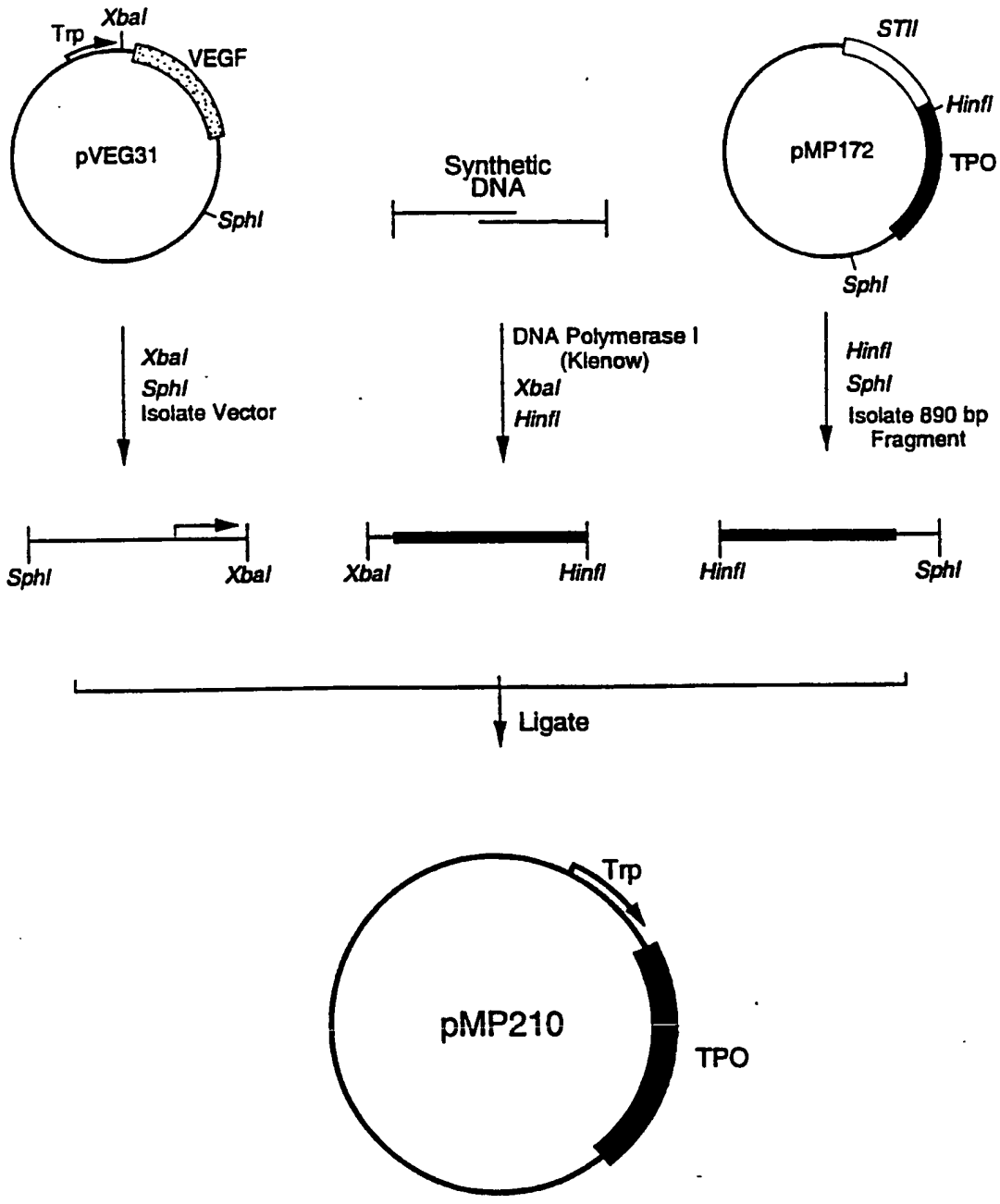

Trp
XbaI
VEGF
pVEG31
SphI
Synthetic DNA
STII
HinfI
pMP172
TPO
SphI
XbaI
SphI
Isolate Vector
DNA Polymerase I (Klenow)
XbaI
HinfI
HinfI
SphI
Isolate 890 bp Fragment
SphI
XbaI
XbaI
HinfI
HinfI
SphI
Ligate
Trp
pMP210
TPO

FIG.38

	Met	Ser	Pro	Ala	Pro	Pro	Ala
MP210 Bank	ATG	TCN	CCN	GCN	CCN	CCN	GCN
MP210-1	ATG	TCT	CCA	GCG	CCG	CCA	GCG
MP210-T8	ATG	TCG	CCT	GCT	CCA	CCT	GCT
MP210-21	ATG	TCG	CCA	GCG	CCA	CCA	GCC
MP210-24	ATG	TCC	CCA	GCC	CCA	CCC	GCA
MP210-25	ATG	TCG	CCA	GCG	CCG	CCA	GCG

FIG.39

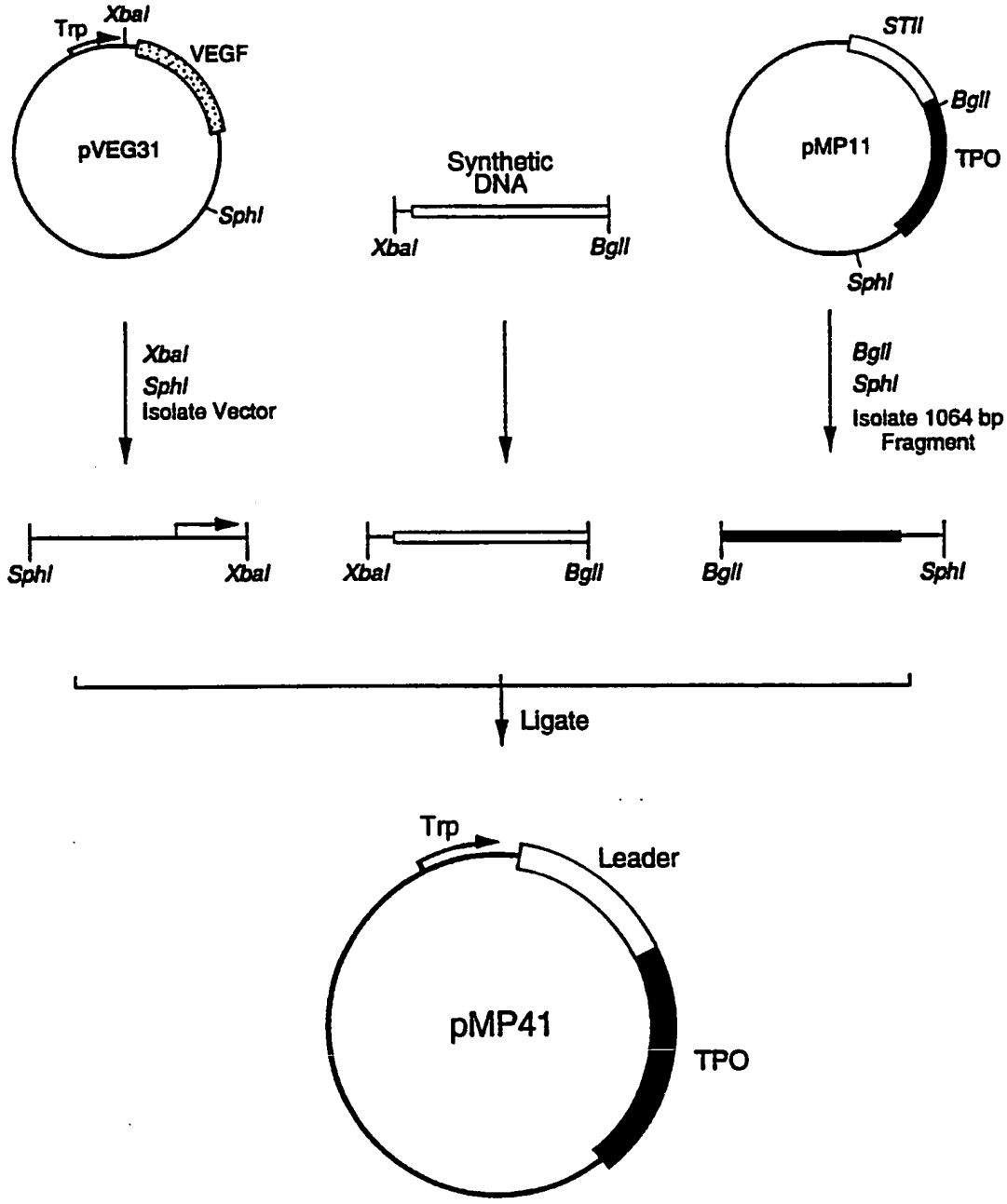

Trp
XbaI
VEGF
pVEG31
SphI
Synthetic DNA
XbaI
BglI
STII
BglI
pMP11
TPO
SphI
XbaI
SphI
Isolate Vector
BglI
SphI
Isolate 1064 bp Fragment
SphI
XbaI
XbaI
BglI
BglI
SphI
Ligate
Trp
Leader
pMP41
TPO

FIG.40

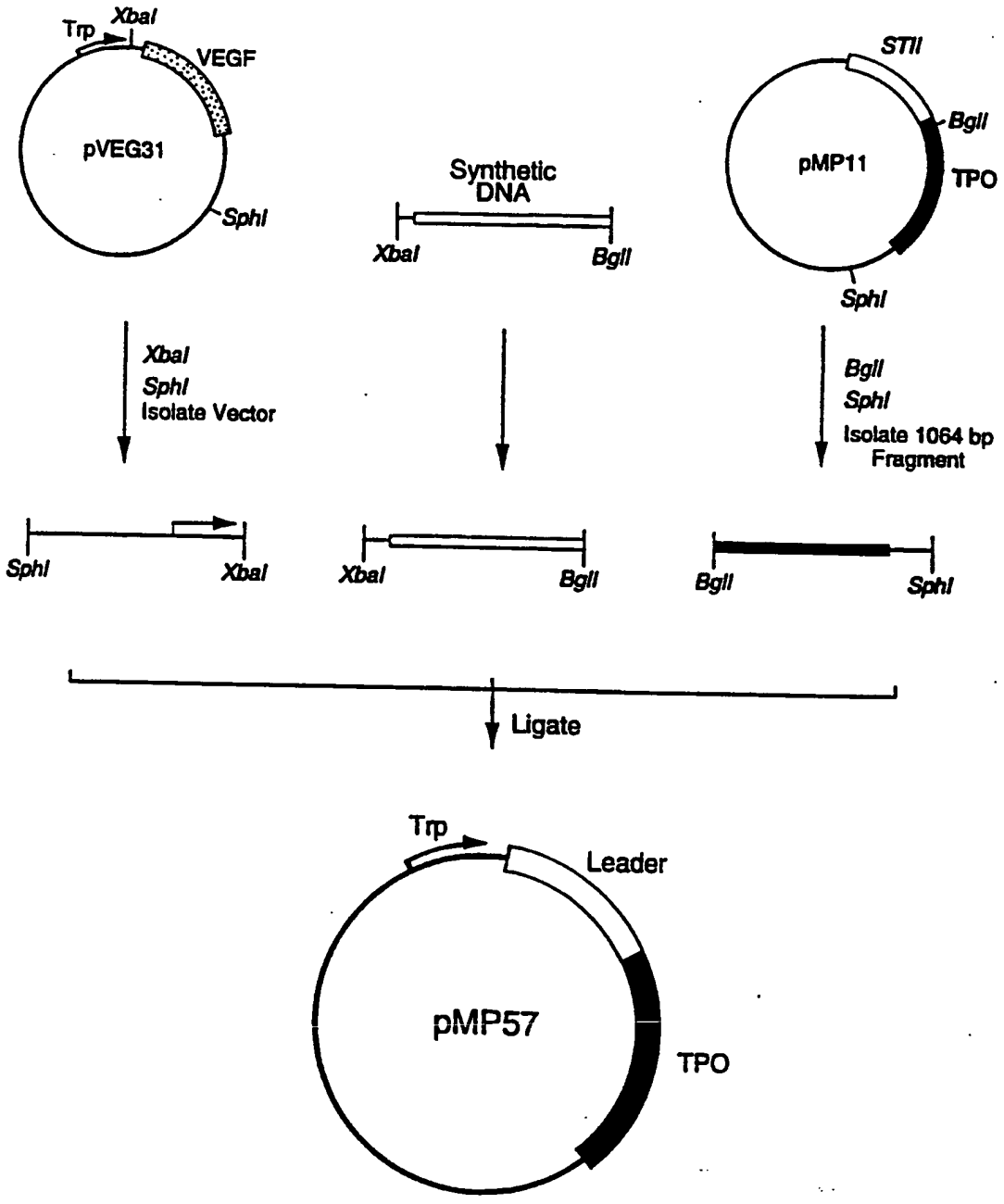

Trp
XbaI
VEGF
pVEG31
SphI
Synthetic DNA
XbaI
BglI
STII
BglI
pMP11
TPO
SphI
XbaI
SphI
Isolate Vector
BglI
SphI
Isolate 1064 bp Fragment
SphI
XbaI
XbaI
BglI
BglI
SphI
Ligate
Trp
Leader
pMP57
TPO

FIG.41

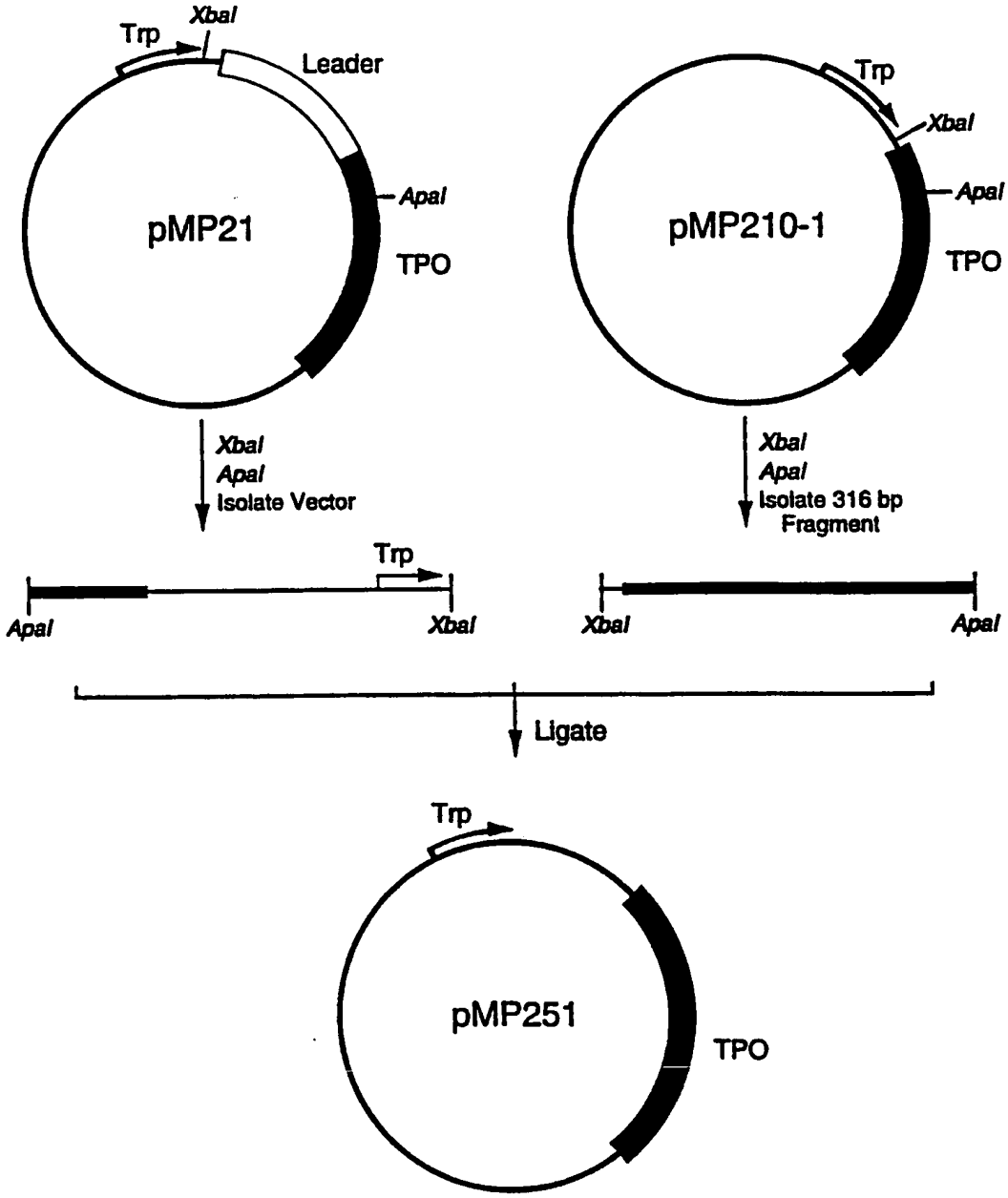

Trp
XbaI
Leader
pMP21
ApaI
TPO
XbaI
ApaI
Isolate Vector
Trp
ApaI
XbaI
Trp
XbaI
pMP210-1
ApaI
TPO
XbaI
ApaI
Isolate 316 bp
Fragment
XbaI
ApaI
Ligate
Trp
pMP251
TPO

FIG.42

INTERNATIONAL SEARCH REPORT

Interna ul Application No
PCT/US 94/14553

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/19 C07K14/52 C07K16/24 A61K38/19

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EXPERIMENTAL HEMATOLOGY, vol. 16,no. 3, March 1988 pages 201-205, MCDONALD 'THROMBOPOETIN:ITS BIOLOGY,PURIFICATION,AND CHARACTERIZATION' see the whole document ---	1-5,12, 36-39
X	EXPERIMENTAL HEMATOLOGY, vol. 17,no. 8, September 1989 pages 865-871, MCDONALD ET AL 'A FOUR-STEP PROCEDURE FOR THE PURIFICATION OF THROMBOPOIETIN' see the whole document --- -/--	1-5,12, 36-39

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 May 1995	07. 06. 95
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer Sitch, W

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	THE JOURNAL OF LABORATORY AND CLINICAL MEDICINE, vol. 106,no. 2, August 1985 pages 162-174, MCDONALD ET AL 'STUDIES ON THE PURIFICATION OF THROMBOPOIETIN FROM KIDNEY CELL CULTURE MEDIUM' see the whole document ---	1-5,12, 36-39
X	EXPERIMENTAL HEMATOLOGY, vol. 2,no. 6, 1974 pages 355-361, MCDONALD ET AL 'PURIFICATION AND ASSAY OF THROMBOPOIETIN' see the whole document ---	1-5,12, 36-39
A	THE EMBO JOURNAL, vol. 12,no. 7, 1993 pages 2645-2653, SKODA ET AL 'MURINE C-MPL:A MEMBER OF THE HEMATOPOIETIC GROWTH FACTOR RECEPTOR SUPERFAMILY THAT TRANSDUCES A PROLIFERATIVE SIGNAL' cited in the application ---	
P,X	NATURE, vol. 369, 16 June 1994 pages 533-538, DE SAUVAGE ET AL 'STIMULATION OF MEGAKARYOCYTOPOIESIS AND THROMBOPOIESIS BY THE C-MPL LIGAND' see the whole document ---	1-40
P,X	CELL, vol. 77, 1 July 1994 pages 1117-1124, BARTLEY ET AL 'IDENTIFICATION AND CLONING OF A MEGAKARYOCYTE GROWTH AND DEVELOPMENT FACTOR THAT IS A LIGAND FOR THE CYTOKINE RECEPTOR MPL' see the whole document ---	1-40
P,X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES,USA, vol. 91,no. 26, 20 December 1994 pages 13023-13027, FOSTER ET AL 'HUMAN THROMBOPOIETIN:GENE STRUCTURE,CDNA SEQUENCE,EXPRESSION,AND CHROMOSOMAL LOCALIZATION' see the whole document --- -/--	1-40

1

INTERNATIONAL SEARCH REPORT

Intern ıal Application No
PCT/US 94/14553

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	FEBS LETTERS, vol. 353,no. 1, 10 October 1994 pages 57-61, SOHMA ET AL 'MOLECULAR CLONING AND CHROMOSOMAL LOCALIZATION OF THE HUMAN THROMBOPOIETIN GENE' see the whole document	1-40

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

THROMBOPOIETIN

FIELD OF THE INVENTION

This invention relates to the isolation, purification and recombinant or chemical synthesis of proteins that influence survival, proliferation, differentiation or maturation of hematopoietic cells, especially platelet progenitor cells. This invention specifically relates to the cloning and expression of nucleic acids encoding a protein ligand capable of binding to and activating *mpl*, a member of the cytokine receptor superfamily. This invention further relates to the use of these proteins alone or in combination with other cytokines to treat immune or hematopoietic disorders including thrombocytopenia.

BACKGROUND OF THE INVENTION

I. The Hematopoietic System

The hematopoietic system produces the mature highly specialized blood cells known to be necessary for survival of all mammals. These mature cells include; erythrocytes, specialized to transport oxygen and carbon dioxide, T- and B-lymphocytes, responsible for cell- and antibody-mediated immune responses, platelets or thrombocytes, specialized to form blood clots, and granulocytes and macrophages, specialized as scavengers and as accessory cells to combat infection. Granulocytes are further subdivided into; neutrophils, eosinophils, basophils and mast cells, specialized cell types having discrete functions. Remarkably, all of these specialized mature blood cells are derived from a single common primitive cell type, referred to as the pluripotent (or totipotent) stem cell, found primarily in bone marrow (Dexter *et al.*, *Ann. Rev. Cell Biol.*, 3:423-441 [1987]).

The mature highly specialized blood cells must be produced in large numbers continuously throughout the life of a mammal. The vast majority of these specialized blood cells are destined to remain functionally active for only a few hours to weeks (Cronkite *et al.*, *Blood Cells*, 2:263-284 [1976]). Thus, continuous renewal of the mature blood cells, the primitive stem cells themselves, as well as any intermediate or lineage-committed progenitor cell lines lying between the primitive and mature cells, is necessary in order to maintain the normal steady state blood cell needs of the mammal.

At the heart of the hematopoietic system lies the pluripotent stem cell(s). These cells are relatively few in number and undergo self-renewal by proliferation to produce daughter stem cells or are transformed, in a series of differentiation steps,

into increasingly mature lineage-restricted progenitor cells, ultimately forming the highly specialized mature blood cell(s).

For example, certain multipotent progenitor cells, referred to as CFC-Mix, derived from stem cells undergo proliferation (self-renewal) and development to produce colonies containing all the different myeloid cells; erythrocytes, neutrophils, megakaryocytes (predecessors of platelets), macrophages, basophils, eosinophils, and mast cells. Other progenitor cells of the lymphoid lineage undergo proliferation and development into T-cells and B-cells.

Additionally, between the CFC-Mix progenitor cells and myeloid cells lie another rank of progenitor cells of intermediate commitment to their progeny. These lineage-restricted progenitor cells are classified on the basis of the progeny they produce. Thus, the known immediate predecessors of the myeloid cells are: erythroid colony-forming units (CFU-E) for erythrocytes, granulocyte/macrophage colony-forming cells (GM-CFC) for neutrophils and macrophages, megakaryocyte colony-forming cells (Meg-CFC) for megakaryocytes, eosinophil colony-forming cells (Eos-CFC) for eosinophils, and basophil colony-forming cells (Bas-CFC) for mast cells. Other intermediate predecessor cells between the pluripotent stem cells and mature blood cells are known (see below) or will likely be discovered having varying degrees of lineage-restriction and self-renewal capacity.

The underlying principal of the normal hematopoietic cell system appears to be decreased capacity for self-renewal as multipotency is lost and lineage-restriction and maturity is acquired. Thus, at one end of the hematopoietic cell spectrum lies the pluripotent stem cell possessing the capacity for self-renewal and differentiation into all the various lineage-specific committed progenitor cells. This capacity is the basis of bone marrow transplant therapy where primitive stem cells repopulate the entire hematopoietic cell system. At the other end of the spectrum lie the highly lineage-restricted progenitors and their progeny which have lost the ability of self-renewal but have acquired mature functional activity.

The proliferation and development of stem cells and lineage-restricted progenitor cells is carefully controlled by a variety of hematopoietic growth factors or cytokines. The role of these growth factors *in vivo* is complex and incompletely understood. Some growth factors, such as interleukin-3 (IL-3), are capable of stimulating both multipotent stem cells as well as committed progenitor cells of several lineages, including for example, megakaryocytes. Other factors such as granulocyte/macrophage colony-stimulating factor (GM-CSF) was initially thought to be restricted in its action to GM-CFC's. Later, however, it was discovered GM-CSF also influenced the proliferation and development of *interalia* megakaryocytes. Thus, IL-3 and GM-CSF were found to have overlapping biological activities, although with

differing potency. More recently, both interleukin-6 (IL-6) and interleukin-11 (IL-11), while having no apparent influence on meg-colony formation alone, act synergistically with IL-3 to stimulate maturation of megakaryocytes (Yonemura *et al.*, *Exp. Hematol.*, **20**:1011-1016 [1992]).

Thus, hematopoietic growth factors may influence growth and differentiation of one or more lineages, may overlap with other growth factors in affecting a single progenitor cell line, or may act synergistically with other factors.

It also appears that hematopoietic growth factors can exhibit their effect at different stages of cell development from the totipotent stem cell through various committed lineage-restricted progenitors to the mature blood cell. For example, erythropoietin (epo) appears to promote proliferation only of mature erythroid progenitor cells. IL-3 appears to exert its effect earlier influencing primitive stem cells and intermediate lineage-restricted progenitor cells. Other growth factors such as stem cell factor (SCF) may influence even more primitive cell development.

It will be appreciated from the foregoing that novel hematopoietic growth factors that affect survival, proliferation, differentiation or maturation of any of the blood cells or predecessors thereof would be useful, especially to assist in the re-establishment of a diminished hematopoietic system caused by disease or after radiation- or chemo-therapy.

II. Megakaryocytopoiesis - Platelet Production

Regulation of megakaryocytopoiesis and platelet production has been reviewed by: Mazur, *Exp. Hematol.*, **15**:248 [1987] and Hoffman, *Blood,* **74**:1196-1212 [1989]. Briefly, bone marrow pluripotent stem cells differentiate into megakaryocytic, erythrocytic, and myelocytic cell lines. It is believed there is a hierarchy of committed megakaryocytic progenitor cells between stem cells and megakaryocytes. At least three classes of megakaryocytic progenitor cells have been identified, namely; burst forming unit megakaryocytes (BFU-MK), colony-forming unit megakaryocytes (CFU-MK), and light density megakaryocyte progenitor cells (LD-CFU-MK). Megakaryocytic maturation itself is a continuum of development that has been separated into stages based on standard morphologic criteria. The earliest recognizable member of the megakaryocyte (MK or meg) family are the megakaryoblasts. These cells are initially 20 to 30 μm in diameter having basophilic cytoplasm and a slightly irregular nucleus with loose, somewhat reticular chromatin and several nucleoli. Later, megakaryoblasts may contain up to 32 nuclei (ploypioid), but the cytoplasm remains sparse and immature. As maturation proceeds, the nucleus becomes more lobulate and pyknotic, the cytoplasm increases in quantity and becomes more acidophilic and granular. The most mature cells of this family may give the

appearance of releasing platelets at their periphery. Normally, less than 10% of megakaryocytes are in the blast stage and more than 50% are mature. Arbitrary morphologic classifications commonly applied to the megakaryocyte series are megakaryoblast for the earliest form; promegakaryocyte or basophilic megakaryocyte for the intermediate form; and mature (acidophilic, granular, or platelet-producing) megakaryocyte for the late forms. The mature megakaryocyte extends filaments of cytoplasm into sinusoidal spaces where they detach and fragment into individual platelets (Williams *et al., Hematology*, 1972).

Megakaryocytopoiesis is believed to involve several regulatory factors (Williams *et al., Br. J. Haematol.*, **52**:173 [1982] and Williams *et al., J. Cell Physiol.*, **110**:101 [1982]). The early level of megakaryocytopoiesis is postulated as being mitotic, concerned with cell proliferation and colony initiation from CFU-MK but is not affected by platelet count (Burstein *et al., J. Cell Physiol.*, **109**:333 [1981] and Kimura *et al., Exp. Hematol.*, **13**:1048 [1985]). The later stage of maturation is non-mitotic, involved with nuclear polyploidization and cytoplasmic maturation and is probably regulated in a feedback mechanism by peripheral platelet number (Odell *et al., Blood*, **48**:765 [1976] and Ebbe *et al., Blood*, **32**:787 [1968]).

The existence of a distinct and specific megakaryocyte colony-stimulating factor (MK-CSF) has been disputed (Mazur, *Exp. Hematol.*, **15**:340-350 [1987]). However most authors believe that a process so vital to survival as platelet production would be regulated by cytokine(s) exclusively responsible for this process. The hypothesis that megakaryocyte/platelet specific cytokine(s) exist has provided the basis for more than 30 years of search - but to date no such cytokine has been purified, sequenced and established by assay as a unique MK-CSF (TPO).

Although it has been reported that MK-CSF's have been partly purified from experimentally produced thrombocytopenia (Hill *et al., Exp. Hematol.*, **14**:752 [1986]) and human embryonic kidney conditioned medium [CM] (McDonald *et al., J. Lab. Clin. Med.*, **85**:59 [1975]) and in man from a plastic anemia and idiopathic thrombocytopenic purpura urinary extracts (Kawakita *et al., Blood*, **6**:556 [1983]) and plasma (Hoffman *et al., J. Clin. Invest.*, **75**:1174 [1985]), their physiological function is as yet unknown in most cases.

The conditioned medium of pokeweed mitogen-activated spleen cells (PWM-SpCM) and the murine myelomonocyte cell line WEHI-3 (WEHI-3CM) have been used as megakaryocyte potentiators. PWM-SpCM contains factors enhancing CFU-MK growth (Metcalf *et al., Pro. Natl. Acad. Sci., USA*, **72**:1744-1748 [1975]; Quesenberry *et al., Blood*, **65**:214 [1985]; and Iscove, N.N., in *Hematopoietic Cell Differentiation, ICN-UCLA Symposia on Molecular and Cellular Biology*, Vol. 10, Golde

et al., eds. [New York, Academy Press] pp 37-52 [1978]), one of which is interleukin-3 (IL-3), a multilineage colony stimulating factor (multi-CSF [Burstein, *Blood Cells*, **11**:469 [1986]). The other factors in this medium have not yet been identified and isolated. WEHI-3 is a murine myelomonocytic cell line secreting relatively large amounts of IL-3 and smaller amounts of GM-CSF. IL-3 has been found to potentiate the growth of a wide range of hematopoietic cells (Ihle *et al.*, *J. Immunol.*, **13**:282 [1983]). IL-3 has also been found to synergize with many of the known hematopoietic hormones or growth factors (Bartelmez *et al.*, *J. Cell Physiol.*, **122**:362-369 [1985] and Warren *et al.*, *Cell*, **46**:667-674 [1988]), including both erythropoietin (EPO) and interleukin-1 (IL-1), in the induction of very early multipotential precursors and the formation of very large mixed hematopoietic colonies.

Other sources of megakaryocyte potentiators have been found in the conditioned media of murine lung, bone, macrophage cell lines, peritoneal exudate cells and human embryonic kidney cells. Despite certain conflicting data (Mazur, *Exp. Hematol.*, **15**:340- 350 [1987]), there is some evidence (Geissler *et al.*, *Br. J. Haematol.*, **60**:233-238 [1985]) that activated T lymphocytes rather than monocytes play an enhancing role in megakaryocytopoiesis. These findings suggest that activated T-lymphocyte secretions such as interleukins may be regulatory factors in MK development (Geissler *et al.*, *Exp. Hematol.*, **15**:845-853 [1987]). A number of studies on megakaryocytopoiesis with purified erythropoietin EPO (Vainchenker *et al.*, *Blood*, **54**:940 [1979]; McLeod *et al.*, *Nature*, **261**:492-4 [1976]; and Williams *et al.*, *Exp. Hematol.*, **12**:734 [1984]) indicate that this hormone has an enhancing effect on MK colony formation. This has also been demonstrated in both serum-free and serum-containing cultures and in the absence of accessory cells (Williams *et al.*, *Exp. Hematol.*, **12**:734 [1984]). EPO was postulated to be involved more in the single and two-cell stage aspects of megakaryocytopoiesis as opposed to the effect of PWM-SpCM which was involved in the four-cell stage of megakaryocyte development. The interaction of all these factors on both early and late phases of megakaryocyte development remains to be elucidated.

Data produced from several laboratories suggests that the only multi-lineage factors that individually have MK-colony stimulating activity are GM-CSF and IL-3 and, to a lesser extent, the B-cell stimulating factor IL-6 (Ikebuchi *et al.*, *Proc. Natl. Acad. Sci. USA*, **84**:9035 [1987]). More recently, several authors have reported that IL-11 and leukemia inhibitory factor (LIF) act synergistically with IL-3 to increase megakaryocyte size and ploidy (Yonemura *et al.*, *British Journal of Hematology*, **84**:16-23 [1993]; Burstein *et al.*, *J. Cell. Physiol.*, **153**:305-312 [1992]; Metcalf *et al.*, *Blood*, **76**:50-56 [1990]; Metcalf *et al.*, *Blood*, **77**:2150-2153 [1991];

Bruno *et al.*, *Exp. Hematol.*, 19:378-381 [1991]; and Yonemura *et al.*, *Exp. Hematol.*, 20:1011-1016 [1992]).

Other documents of interest include: Eppstein *et al.*, U.S. Patent No. 4,962,091; Chong, U.S. Patent No. 4,879,111; Fernandes *et al.*, U.S. Patent No. 4,604,377; Wissler *et al.*, U.S. Patent No. 4,512,971; Gottlieb, U.S. Patent No. 4,468,379; Bennett *et al.*, U.S. Patent No. 5,215,895; Kogan *et al.*, U.S. Patent No. 5,250,732; Kimura *et al.*, *Eur. J. Immunol.*, 20(9):1927-1931 [1990]; Secor *et al.*, *J. of Immunol.*, 144(4):1484-1489 [1990]; Warren *et al.*, *J. of Immunol.*, 140(1):94-99 [1988]; Warren *et al.*, *Exp. Hematol.*, 17(11):1095-1099 [1989]; Bruno *et al.*, *Exp. Hematol.*, 17(10):1038-1043 [1989]; Tanikawa *et al.*, *Exp. Hematol.*, 17(8):883-888 [1989]; Koike *et al.*, *Blood*, 75(12):2286-2291 [1990]; Lotem, *Blood*, 75(5):1545-1551 [1989]; Rennick *et al.*, *Blood*, 73(7):1828-1835 [1989]; and Clutterbuck *et al.*, *Blood*, 73(6):1504-1512 [1989].

III. Thrombocytopenia

Platelets are critical elements of the blood clotting mechanism. Depletion of the circulating level of platelets, called thrombocytopenia, occurs in various clinical conditions and disorders. Thrombocytopenia is commonly defined as a platelet count below 150 X 10^9 per liter. The major causes of thrombocytopenia can be broadly divided into three categories on the basis of platelet life span, namely; (1) impaired production of platelets by the bone marrow, (2) platelet sequestration in the spleen (splenomegaly), or (3) increased destruction of platelets in the peripheral circulation (*e.g.*, autoimmune thrombocytopenia or chemo- and radiation-therapy). Additionally, in patients receiving large volumes of rapidly administered platelet-poor blood products, thrombocytopenia may develop due to dilution.

The clinical bleeding manifestations of thrombocytopenia depend on the severity of thrombocytopenia, its cause, and possible associated coagulation defects. In general, patients with platelet counts between 20 and 100 X 10^9 per liter are at risk of excessive post traumatic bleeding, while those with platelet counts below 20 X 10^9 per liter may bleed spontaneously. These latter patients are candidates for platelet transfusion with attendant immune and viral risk. For any given degree of thrombocytopenia, bleeding tends to be more severe when the cause is decreased production rather than increased destruction of platelets. In the latter situation, accelerated platelet turnover results in the circulation of younger, larger and hemostatically more effective platelets. Thrombocytopenia may result from a variety of disorders briefly described below. A more detailed description may be found in Schafner, A. I., "Thrombocytopenia and Disorders of Platelet Function," *Internal*

Medicine, 3rd Ed., John J. Hutton *et al.*, Eds., Little Brown and Co., Boston/Toronto/London [1990].

(a) Thrombocytopenia due to impaired platelet production

Causes of congenital thrombocytopenia include constitutional aplastic anemia (Fanconi syndrome) and congenital amegakaryocytic thrombocytopenia, which may be associated with skeletal malformations. Acquired disorders of platelet production are caused by either hypoplasia of megakaryocytes or ineffective thrombopoiesis. Megakaryocytic hypoplasia can result from a variety of conditions, including marrow aplasia (including idiopathic forms or myelosuppression by chemotherapeutic agents or radiation therapy), myelfibrosis, leukemia, and invasion of the bone marrow by metastatic tumor or granulomas. In some situations, toxins, infectious agents, or drugs may interfere with thrombopoiesis relatively selectively; examples include transient thrombocytopenias caused by alcohol and certain viral infections and mild thrombocytopenia associated with the administration of thiazide diuretics. Finally, ineffective thrombopoiesis secondary to megaloblastic processes (folate or B_{12} deficiency) can also cause thrombocytopenia, usually with coexisting anemia and leukopenia.

Current treatment of thrombocytopenias due to decreased platelet production depends on identification and reversal of the underlying cause of the bone marrow failure. Platelet transfusions are usually reserved for patients with serious bleeding complications, or for coverage during surgical procedures, since isoimmunization may lead to refractoriness to further platelet transfusions. Mucosal bleeding resulting from severe thrombocytopenia may be ameliorated by the oral or intravenous administration of the antifibrinolytic agents. Thrombotic complications may develop, however, if antifibrinolytic agents are used in patients with disseminated intravascular coagulation (DIC).

(b) Thrombocytopenia due to splenic sequestration

Splenomegaly due to any cause may be associated with mild to moderate thrombocytopenia. This is a largely passive process (hypersplenism) of splenic platelet sequestration, in contrast to the active destruction of platelets by the spleen in cases of immunomediated thrombocytopenia discussed below. Although the most common cause of hypersplenism is congestive splenomegaly from portal hypertension due to alcoholic cirrhosis, other forms of congestive, infiltrative, or lymphoproliferative splenomegaly are also associated with thrombocytopenia. Platelet counts generally do not fall below 50×10^9 per liter as a result of hypersplenism alone.

(c) Thrombocytopenia due to nonimmune-mediated platelet destruction

Thrombocytopenia can result from the accelerated destruction of platelets by various nonimmunologic processes. Disorders of this type include disseminated intravascular coagulation, prosthetic intravascular devices, extra corporeal circulation of the blood, and thrombotic microangiopathies such as thrombotic thrombocytic purpura. In all of these situations, circulating platelets that are exposed to either artificial surfaces or abnormal vascular intima either are consumed at these sites or are damaged and then prematurely cleared by the reticuloendothelial system. Disease states or disorders in which disseminated intravascular coagulation (DIC) may arise are set forth in greater detail in Braunwald *et al.* (eds), *Harrison's Principles of Internal Medicine*, 11th Ed., p.1478, McGraw Hill [1987]. Intravascular prosthetic devices, including cardiac valves and intra-aortic balloons can cause a mild to moderate destructive thrombocytopenia and transient thrombocytopenia in patients undergoing cardiopulmonary bypass or hemodialysis may result from consumption or damage of platelets in the extra corporeal circuit.

(d) Drug-induced immune thrombocytopenia

More than 100 drugs have been implicated in immunologically mediated thrombocytopenia. However, only quinidine, quinine, gold, sulfonamides, cephalothin, and heparin have been well characterized. Drug-induced thrombocytopenia is frequently very severe and typically occurs precipitously within days while patients are taking the sensitizing medication.

(e) Immune (autoimmune) thrombocytopenic purpura (ITP)

ITP in adults is a chronic disease characterized by autoimmune platelet destruction. The autoantibody is usually IgG although other immunoglobulins have also been reported. Although the autoantibody of ITP has been found to be associated with platelet membrane $GPIIb_bIII_a$, the platelet antigen specificity has not been identified in most cases. Extravascular destruction of sensitized platelets occurs in the reticuloendothelial system of the spleen and liver. Although over one-half of all cases of ITP are idiopathic, many patients have underlying rheumatic or autoimmune diseases (*e.g.*, systemic lupus erythematosus) or lymphoproliferative disorders (*e.g.*, chronic lymphocytic leukemia).

(f) HIV-Induced ITP

ITP is an increasingly common complication of HIV infection (Morris *et al.*, *Ann. Intern. Med.*, 96:714-717 [1982]), and can occur at any stage of the disease progression, both in patients diagnosed with the Acquired Immune Deficiency Syndrome (AIDS), those with AIDS-related complex, and those with HIV infection but without AIDS symptoms. HIV infection is a transmissible disease ultimately characterized by a profound deficiency of cellular immune function as well as the

occurrence of opportunistic infection and malignancy. The primary immunologic abnormality resulting from infection by HIV is the progressive depletion and functional impairment of T lymphocytes expressing the CD4 cell surface glycoprotein (Lane *et al., Ann. Rev. Immunol.*, 3:477 [1985]). The loss of CD4 helper/inducer T cell function probably underlies the profound defects in cellular and humoral immunity leading to the opportunistic infections and malignancies characteristic of AIDS (H. Lane *supra*).

Although the mechanism of HIV-associated ITP is unknown, it is believed to be different from the mechanism of ITP not associated with HIV infection. (Walsh *et al., N. Eng. J. Med.*, 311:635-639 [1984]; and Ratner, *Am. J. Med.*, 86:194-198 [1989]).

IV. Current Therapy for Thrombocytopenia

The therapeutic approach to the treatment of patients with thrombocytopenia is dictated by the severity and urgency of the clinical situation. The treatment is similar for HIV-associated and non-HIV-related thrombocytopenia, and although a number of different therapeutic approaches have been used, the therapy remains controversial.

Platelet counts in patients diagnosed with thrombocytopenia have been successfully increased by glucocorticoid (*e.g.*, prednisolone) therapy, however in most patients, the response is incomplete, or relapse occurs when the glucocorticoid dose is reduced or its administration is discontinued. Based upon studies with patients having HIV-associated ITP, some investigators have suggested that glucocorticold therapy may result in predisposition to AIDS. Glucocorticoids are usually administered if platelet count falls below 20 X 10^9/liter or when spontaneous bleeding occurs.

For patients refractory to glucocorticoids, the compound:

4-(2-chlorphenyl)-9-methyl-2-[3-(4-morpholinyl)-3-propanon-1-yl]6H-thieno[3,2,f][1,2,4]triazolo[4,3,a,][1,4]diazepin (WEB 2086)

has been successfully used to treat a severe case of non HIV-associated ITP. A patient having platelet counts of 37,000-58,000/µl was treated with WEB 2086 and after 1-2 weeks treatment platelet counts increased to 140,000-190,000/µl. (EP 361,077 and Lohman *et al., Lancet*, 1147 [1988]).

Although the optimal treatment for acquired amegakaryocytic thrombocytopenia purpura (AATP) is uncertain, antithymocyte globulin (ATG), a horse antiserum to human thymus tissue, has been shown to produce prolonged complete remission (Trimble *et al., Am. J. Hematol.*, 37:126-127 [1991]). A recent report however, indicates that the hematopoietic effects of ATG are attributable to thimerosal, where presumably the protein acts as a mercury carrier (Panella *et al., Cancer Research*, 50:4429-4435 [1990]).

Good results have been reported with splenectomy. Splenectomy removes the major site of platelet destruction and a major source of autoantibody production in many patients. This procedure results in prolonged treatment-free remissions in a large number of patients. However, since surgical procedures are generally to be avoided in immune compromised patients, splenectomy is recommended only in severe cases of thrombocytopenia (*e.g.* severe HIV-associated ITP), in patients who fail to respond to 2 to 3 weeks of glucocorticoid treatment, or do not achieve sustained response after discontinuation of glucocorticoid administration. Based upon current scientific knowledge, it is unclear whether splenectomy predisposes patients to AIDS.

In addition to prednisolone therapy and splenectomy, certain cytotoxic agents, *e.g.*, vincristine, and azidothimidine (AZT, zidovudine) also show promise in treating HIV-induced ITP; however, the results are preliminary.

It will be appreciated from the foregoing that one way to treat thrombocytopenia would be to obtain an agent capable of accelerating the differentiation and maturation of megakaryocytes or precursors thereof into the platelet-producing form. Considerable efforts have been expended on identifying such an agent, commonly referred to as "thrombopoietin" (TPO). Other names for TPO commonly found in the literature include; thrombocytopoiesis stimulating factor (TSF), megakaryocyte colony-stimulating factor (MK-CSF), megakaryocyte-stimulating factor and megakaryocyte potentiator. TPO activity was observed as early as 1959 (Rak *et al., Med. Exp.,* 1:125) and attempts to characterize and purify this agent have continued to the present day. While reports of partial purification of TPO-active polypeptides exist (see, for example, Tayrien *et al., J. Biol. Chem.*, **262**:3262 [1987] and Hoffman *et al., J. Clin. Invest.* **75**:1174 [1985]), others have postulated that TPO is not a discrete entity in its own right but rather is simply the polyfunctional manifestation of a known hormone (IL- 3, Sparrow *et al., Prog. Clin. Biol. Res.,* **215**:123 [1986]). Regardless of its form or origin, a molecule possessing thrombopoietic activity would be of significant therapeutic value. Although no protein has been unambiguously identified as TPO, considerable interest surrounds the recent discovery that *mpl,* a putative cytokine receptor, may transduce a thrombopoietic signal.

V. *Mpl* is a Megakaryocytopoietic Cytokine Receptor

It is believed that the proliferation and maturation of hematopoietic cells is tightly regulated by factors that positively or negatively modulate pluripotential stem cell proliferation and multilineage differentiation. These effects are mediated through the high-affinity binding of extracellular protein factors to specific cell surface receptors. These cell surface receptors share considerable homology and are generally

classified as members of the cytokine receptor superfamily. Members of the superfamily include receptors for: IL-2 (β and γ chains) (Hatakeyama *et al.*, *Science*, **244**:551-556 [1989]; Takeshita *et al.*, *Science*, **257**:379-382 [1991]); IL-3 (Itoh *et al.*, *Science*, **247**:324-328 [1990]; Gorman *et al.*, *Proc. Natl. Acad. Sci. USA*, **87**:5459-5463 [1990]; Kitamura *et al.*, *Cell*, **66**:1165-1174 [1991a]; Kitamura *et al.*, *Proc. Natl. Acad. Sci. USA*, **88**:5082-5086 [1991b]), IL-4 (Mosley *et al.*, *Cell*, **59**:335-348 [1989], IL-5 (Takaki *et al.*, *EMBO J.*, **9**:4367-4374 [1990]; Tavernier *et al.*, *Cell*, **66**:1175-1184 [1991]), IL-6 (Yamasaki *et al.*, *Science*, **241**:825-828 [1988]; Hibi *et al.*, *Cell*, **63**:1149-1157 [1990]), IL-7 (Goodwin *et al.*, *Cell*, **60**:941-951 [1990]), IL-9 (Renault *et al.*, *Proc. Natl. Acad. Sci. USA*, **89**:5690-5694 [1992]), granulocyte-macrophage colony-stimulating factor (GM-CSF) (Gearing *et al.*, *EMBO J.*, **8**:3667-3676 [1991]; Hayashida *et al.*, *Proc. Natl. Acad. Sci. USA*, **244**:9655-9659 [1990]), granulocyte colony-stimulating factor (G-CSF) (Fukunaga *et al.*, *Cell*, **61**:341-350 [1990a]; Fukunaga *et al.*, *Proc. Natl. Acad. Sci. USA*, **87**:8702-8706 [1990b]; Larsen *et al.*, *J. Exp. Med.*, **172**:1559-1570 [1990]), EPO (D'Andrea *et al.*, *Cell*, **57**:277-285 [1989]; Jones *et al.*, *Blood*, **76**:31-35 [1990]), Leukemia inhibitory factor (LIF) (Gearing *et al.*, *EMBO J.*, **10**:2839-2848 [1991]), oncostatin M (OSM) (Rose *et al.*, *Proc. Natl. Acad. Sci. USA*, **88**:8641-8645 [1991]) and also receptors for prolactin (Boutin *et al.*, *Proc. Natl. Acad. Scl. USA*, **88**:7744-7748 [1988]; Edery *et al.*, *Proc. Natl. Acad. Sci. USA*, **86**:2112-2116 [1989]), growth hormone (GH) (Leung *et al.*, *Nature*, **330**:537-543 [1987]) and ciliary neurotrophic factor (CNTF) (Davis *et al.*, *Science*, **253**:59-63 [1991].

Members of the cytokine receptor superfamily may be grouped into three functional categories (for review see Nicola *et al.*, *Cell*, **67**:1-4 [1991]). The first class comprises single chain receptors, such as erythropoietin receptor (EPO-R) or granulocyte colony stimulating factor receptor (G-CSF-R), which bind ligand with high affinity via the extracellular domain and also generate an intracellular signal. A second class of receptors, so called α-subunits, includes interleukin-6 receptor (IL6-R), granulocyte-macrophage colony stimulating factor receptor (GM-CSF-R), interleukin-3 receptor (IL3-Rα) and other members of the cytokine receptor superfamily. These α-subunits bind ligand with low affinity but cannot transduce an intracellular signal. A high affinity receptor capable of signaling is generated by a heterodimer between an α-subunit and a member of a third class of cytokine receptors, termed β-subunits, *e.g.*, β_c, the common β-subunit for the three α-subunits IL3-Rα and GM-CSF-R.

Evidence that *mpl* is a member of the cytokine receptor superfamily comes from sequence homology (Gearing, *EMBO J.*, **8**:3667-3676 [1988]; Bazan, *Proc.*

Natl. Acad. Sci. USA, **87**:6834-6938 [1990]; Davis *et al., Science*, **253**:59-63 [1991] and Vigon *et al., Proc. Natl. Acad. Sci. USA*, **89**:5640-5644 [1992]) and its ability to transduce proliferative signals.

Deduced protein sequence from molecular cloning of murine c-*mpl* reveals this protein is homologous to other cytokine receptors. The extracellular domain contains 465 amino acid residues and is composed of two subdomains each with four highly conserved cysteines and a particular motif in the N-terminal subdomain and in the C-terminal subdomain. The ligand-binding extracellular domains are predicted to have similar double β-barrel fold structural geometries. This duplicated extracellular domain is highly homologous to the signal transducing chain common to IL-3, IL-5 and GM-CSF receptors as well as the low-affinity binding domain of LIF (Vigon *et al.*, *Oncogene*, **8**:2607-2615 [1993]). Thus *mpl* may belong to the low affinity ligand binding class of cytokine receptors.

A comparison of murine *mpl* and mature human *mpl* P, reveals these two proteins show 81% sequence identity. More specifically, the N-terminus and C-terminus extracellular subdomains share 75% and 80% sequence identity respectively. The most conserved *mpl* region is the cytoplasmic domain showing 91% amino acid identity, with a sequence of 37 residues near the transmembrane domain being identical in both species. Accordingly, *mpl* is reported to be one of the most conserved members of the cytokine receptor superfamily (Vigon *supra*).

Evidence that *mpl* is a functional receptor capable of transducing a proliferative signal comes from construction of chimeric receptors containing an extracellular domain from a cytokine receptor having high affinity for a known cytokine with the *mpl* cytoplasmic domain. Since no known ligand for *mpl* has been reported, it was necessary to construct the chimeric high affinity ligand binding extracellular domain from a class one cytokine receptor such as IL-4R or G-CSFR. Vigon *et al.*, *supra* fused the extracellular domain of G-CSFR with both the transmembrane and cytoplasmic domain of c-*mpl*. An IL-3 dependent cell line, BAF/B03 (Ba/F3) was transfected with the G-CSFR/*mpl* chimera along with a full length G-CSFR control. Cells transfected with the chimera grew equally well in the presence of cytokine IL-3 or G-CSF. Similarly, cells transfected with G-CSFR also grew well in either IL-3 or G-CSF. All cells died in the absence of growth factors. A similar experiment was conducted by Skoda *et al.*, *EMBO J.*, **12**(7):2645-2653 [1993] in which both the extracellular and transmembrane domains of human IL-4 receptor (hIL-4-R) were fused to the murine *mpl* cytoplasmic domain, and transfected into a murine IL-3 dependent Ba/F3 cell line. Ba/F3 cells transfected with wild type hIL-4-R proliferated normally in the presence of either of the species specific IL-4 or IL-3. Ba/F3 cells transfected with hIL-4R/*mpl* proliferated

normally in the presence of hIL-4 (in the presence or absence of IL-3) demonstrating that in Ba/F3 cells the *mpl* cytoplasmic domain contains all the elements necessary to transduce a proliferative signal.

These chimeric experiments demonstrate the proliferation signaling capability of the *mpl* cytoplasmic domain but are silent regarding whether the *mpl* extracellular domain can bind a ligand. These results are consistent with at least two possibilities, namely, *mpl* is a single chain (class one) receptor like EPO-R or G-CSFR or it is a signal transducing β-subunit (class three) requiring an α-subunit like IL-3 (Skoda *et al. supra*).

VI. *Mpl* Ligand Is a Thrombopoietin (TPO)

As described above, it has been suggested that serum contains a unique factor, sometimes referred to as thrombopoietin (TPO), that acts synergistically with various other cytokines to promote growth and maturation of megakaryocytes. No such natural factor has ever been isolated from serum or any other source even though considerable effort has been expended by numerous groups. Even though it is not known whether *mpl* is capable of directly binding a megakaryocyte stimulating factor, recent experiments demonstrate that *mpl* is involved in proliferative signal transduction from a factor or factors found in the serum of patients with aplastic bone marrow (Methia *et al.*, *Blood*, 82(5):1395-1401 [1993]).

Evidence that a unique serum colony-forming factor distinct from IL-1α, IL-3, IL-4, IL-6, IL-11, SCF, EPO, G-CSF, and GM-CSF transduces a proliferative signal through *mpl* comes from examination of the distribution of c-*mpl* expression in primitive and committed hematopoietic cell lines and from *mpl* antisense studies in one of these cell lines.

Using reverse transcriptase (RT)-PCR in immuno-purified human hematopoietic cells, Methia *et al., supra* demonstrated that strong *mpl* mRNA messages were only found in $CD34^+$ purified cells, megakaryocytes and platelets. $CD34^+$ cells purified from bone marrow (BM) represents about 1% of all BM cells and are enriched in primitive and committed progenitors of all lineages (*e.g.*, erythroid, granulomacrophage, and megakaryocytic).

Mpl antisense oligodeoxynucleotides were shown to suppress megakaryocytic colony formation from the pluripotent $CD34^+$ cells cultured in serum from patients with aplastic marrow (a rich source of megakaryocyte colony-stimulating activity [MK-CSA]). These same antisense oligodeoxynucleotides had no effect on erythroid or granulomacrophage colony formation.

Whether *mpl* directly bound a ligand and whether the serum factor shown to cause megakaryocytopoiesis acted through *mpl* was still unknown. It had been

suggested, however, that if *mpl* did directly bind a ligand, its amino acid sequence was likely to be highly conserved and have species cross-reactivity owing to the considerable sequence identity between human and murine *mpl* extracellular domains (Vigon *et al., supra* [1993]).

VII. Objects

In view of the foregoing, it will be appreciated there is a current and continuing need in the art to isolate and identify molecules capable of stimulating proliferation, differentiation and maturation of hematopoietic cells, especially megakaryocytes or their predecessors for therapeutic use in the treatment of thrombocytopenia. It is believed such a molecule is a *mpl* ligand and thus there exists a further need to isolate such ligand(s) to evaluate their role(s) in cell growth and differentiation.

Accordingly, it is an object of this invention to obtain a pharmaceutically pure molecule capable of stimulating proliferation, differentiation and/or maturation of megakaryocytes into the mature platelet-producing form.

It is another object to provide the molecule in a form for therapeutic use in the treatment of a hematopoietic disorder, especially thrombocytopenia.

It is a further object of the present invention to isolate, purify and specifically identify protein ligands capable of binding *in vivo* a cytokine superfamily receptor known as *mpl* and to transduce a proliferative signal.

It is still another object to provide nucleic acid molecules encoding such protein ligands and to use these nucleic acid molecules to produce *mpl* binding ligands in recombinant cell culture for diagnostic and therapeutic use.

It is yet another object to provide derivatives and modified forms of the protein ligands including amino acid sequence variants, variant glycoprotein forms and covalent derivatives thereof.

It is an additional object to provide fusion polypeptide forms combining a *mpl* ligand and a heterologous protein and covalent derivatives thereof.

It is still an additional object to provide variant polypeptide forms combining a *mpl* ligand with amino acid additions and substitutions from the EPO sequence to produce a protein capable of regulating proliferation and growth of both platelets and red blood cell progenitors.

It is yet an additional object to prepare immunogens for raising antibodies against *mpl* ligands or fusion forms thereof, as well as to obtain antibodies capable of binding such ligands.

These and other objects of the invention will be apparent to the ordinary artisan upon consideration of the specification as a whole.

SUMMARY OF THE INVENTION

The objects of the invention are achieved by providing an isolated mammalian megakaryocytopoietic proliferation and maturation promoting protein, denominated the "*mpl* ligand" (ML) or "thrombopoietin" (TPO), capable of stimulating proliferation, maturation and/or differentiation of megakaryocytes into the mature platelet-producing form.

This substantially homogeneous protein may be purified from a natural source by a method comprising; (1) contacting a source plasma containing the *mpl* ligand molecules to be purified with an immobilized receptor polypeptide, specifically *mpl* or a *mpl* fusion polypeptide immobilized on a support, under conditions whereby the *mpl* ligand molecules to be purified are selectively adsorbed onto the immobilized receptor polypeptide, (2) washing the immobilized receptor polypeptide and its support to remove non-adsorbed material, and (3) eluting the *mpl* ligand molecules from the immobilized receptor polypeptide to which they are adsorbed with an elution buffer. Preferably the natural source is mammalian plasma or urine containing the *mpl* ligand. Optionally the mammal is aplastic and the immobilized receptor is a *mpl*-IgG fusion.

Optionally, the prefered megakaryocytopoietic proliferation and maturation promoting protein is an isolated substantially homogeneous *mpl* ligand polypeptide made by synthetic or recombinant means.

The "*mpl* ligand" polypeptide or "TPO" of this invention preferably has at least 70% overall sequence identity with the amino acid sequence of the highly purified substantially homogeneous porcine *mpl* ligand polypeptide and at least 80% sequence identity with the "EPO-domain" of the porcine *mpl* ligand polypeptide. Optionally, the *mpl* ligand of this invention is mature human *mpl* ligand (hML), having the mature amino acid sequence provided in **Fig. 1** (SEQ ID NO: 1), or a variant or posttranscriptionally modified form thereof or a protein having about 80% sequence identity with mature human *mpl* ligand. Optionally the *mpl* ligand variant is a fragment, especially an amino-terminus or "EPO-domain" fragment, of the mature human *mpl* ligand (hML). Preferably the amino terminus fragment retains substantially all of the human ML sequence between the first and forth cysteine residues but may contain substantial additions, deletions or substitutions outside that region. According to this embodiment, the fragment polypeptide may be represented by the formula:

X-hML(7-151)-Y

Where hML(7-151) represents the human TPO (hML) amino acid sequence from Cys^{7} through Cys^{151} inclusive; X represents the amino group of Cys^{7} or one or more of the

amino-terminus amino acid residue(s) of the mature hML or amino acid residue extensions thereto such as Met, Tyr or leader sequences containing, for example, proteolytic cleavage sites (*e.g.* Factor Xa or thrombin); and Y represents the carboxy terminal group of Cys^{151} or one or more carboxy-terminus amino acid residue(s) of the mature hML or extensions thereto.

Optionally the *mpl* ligand polypeptide or fragment thereof may be fused to a heterologous polypeptide (chimera). A preferred heterologous polypeptide is a cytokine, colony stimulating factor or interleukin or fragment thereof, especially kit-ligand (KL), IL-1, IL-3, IL-6, IL-11, EPO, GM-CSF or LIF. An optional preferred heterologous polypeptide is an immunoglobin chain, especially human IgG1, IgG2, IgG3, IgG4, IgA, IgE, IgD, IgM or fragment thereof, especially comprising the constant domain of an IgG heavy chain.

Another aspect of this invention provides a composition comprising an isolated *mpl* agonist that is biologically active and is preferably capable of stimulating the incorporation of labeled nucleotides (*e.g.*, ^{3}H-thymidine) into the DNA of IL-3 dependent Ba/F3 cells transfected with human *mpl.* Optionally the *mpl* agonist is biologically active *mpl* ligand and is preferably capable of stimulating the incorporation of ^{35}S into circulating platelets in a mouse platelet rebound assay. Suitable *mpl* agonist include hML_{153}, hML(R153A, R154A), hML2, hML3, hML4, mML, mML2, mML3, pML, and pML2 or fragments thereof.

In another embodiment, this invention provides an isolated antibody capable of binding to the *mpl* ligand. The isolated antibody capable of binding to the *mpl* ligand may optionally be fused to a second polypeptide and the antibody or fusion thereof may be used to isolate and purify *mpl* ligand from a source as described above for immobilized *mpl.* In a further aspect of this embodiment, the invention provides a method for detecting the *mpl* ligand *in vitro* or *in vivo* comprising contacting the antibody with a sample, especially a serum sample, suspected of containing the ligand and detecting if binding has occurred.

In still further embodiments, the invention provides an isolated nucleic acid molecule, encoding the *mpl* ligand or fragments thereof, which nucleic acid molecule may optionally be labeled with a detectable moiety, and a nucleic acid molecule having a sequence that is complementary to, or hybridizes under moderate to highly stringent conditions with, a nucleic acid molecule having a sequence encoding a *mpl* ligand. Preferred nucleic acid molecules are those encoding human, porcine, and murine *mpl* ligand, and include RNA and DNA, both genomic and cDNA. In a further aspect of this embodiment, the nucleic acid molecule is DNA encoding the *mpl* ligand and further comprises a replicable vector in which the DNA is operably linked to control sequences

recognized by a host transformed with the vector. Optionally the DNA is cDNA having the sequence provided in **Fig. 1** 5'-3' (SEQ ID NO: 2), 3'-5' or a fragment thereof. This aspect further includes host cells, preferably CHO cells, transformed with the vector and a method of using the DNA to effect production of *mpl* ligand, preferably comprising expressing the cDNA encoding the *mpl* ligand in a culture of the transformed host cells and recovering the *mpl* ligand from the host cells or the host cell culture. The *mpl* ligand prepared in this manner is preferably human *mpl* ligand.

The invention further includes a method for treating a mammal having a hematopoietic disorder, especially thrombocytopenia, comprising administering a therapeutically effective amount of a *mpl* ligand to the mammal. Optionally the *mpl* ligand is administered in combination with a cytokine, especially a colony stimulating factor or interleukin. Preferred colony stimulating factors or interleukins include; kit-ligand (KL), LIF, G-CSF, GM-CSF, M-CSF, EPO, IL-1, IL-3, IL-6, and IL-11.

The invention further includes a process for isolating and purifying TPO (ML) from a TPO producing microorganism comprising:

(1) disrupting or lysing cells containing TPO,

(2) optionally seperating soluble material from insoluble material containing TPO,

(3) solublizing TPO in the insoluble material with a solublizing buffer,

(4) seperating solublized TPO from other soluble and insoluble material,

(5) refolding TPO in a redox buffer, and

(6) separating properly folded TPO from misfolded TPO.

The process provides for solubilizing the insoluble material containing TPO with a chaotropic agent where the chaotropic agent is selected from a salt of guanidine, sodium thiocyanate, or urea. The process further provides that solublized TPO is seperated from other soluble and insoluble material by one or more steps selected from centrafugation, gel filtration and reverse phase chromotography. The refolding step of the process provides for a redox buffer containing both an oxidizing and reducing agent. Generally, the oxidizing agent is oxygen or a compound containing at least one disulfide bond and the reducing agent is a compound containing at least one free sulfhydryl. Preferably, the oxidizing agent is selected from oxidized glutathione(GSSG) and cystine and the reducing agent is selected from reduced glutathione(GSH) and cysteine. Most preferably the oxidizing agent is oxidized glutathione(GSSG) and the reducing agent is reduced glutathione(GSH). It is also prefered that the molar ratio of the oxidizing agent is equal to or greater then that of the reducing agent. The redox buffer additionally contains a detergent, preferably selected from CHAPS and CHAPSO, present at a level of at least1%. The redox buffer additionally contains NaCl preferably at a concentration range of about 0.1-0.5M, and

glycerol preferably at a concentration greater than 15%. The pH of the redox buffer preferably ranges from about pH 7.5-pH 9.0. and the refolding step is conducted at 4 degrees for 12-48hr. The refolding step produces biologically active TPO in which a disulfide bond is formed between the Cys nearest the amino-terminus with the Cys nearest the carboxy-terminus of the EPO domain.

The invention further includes a process for purifying biologically active TPO from a microorganism comprising:

(1) lysing at least the extracellular membrane of the microorganism,

(2) treating the lysate containing TPO with a chaotropic agent,

(3) refolding the TPO, and

(4) separating impurities and misfolded TPO from properly folded TPO.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 shows the deduced amino acid sequence (SEQ ID NO: 1) of human *mpl* ligand (hML) cDNA and the coding nucleotide sequence (SEQ ID NO: 2). Nucleotides are numbered at the beginning of each line. The 5' and 3' untranslated regions are indicated in lower case letters. Amino acid residues are numbered above the sequence starting at Ser 1 of the mature *mpl* ligand (ML) protein sequence. The boundries of presumed exon 3 are indicated by the arrows and the potential N-glycosylation sites are boxed. Cysteine residues are indicated by a dot above the sequence. The underlined sequence corresponds to the N-terminal sequence determined from *mpl* ligand purified from porcine plasma.

Fig. 2 shows the procedure used for the *mpl* ligand ^{3}H-thymidine incorporation assay. To determine the presence of *mpl* ligand from various sources, the *mpl* P Ba/F3 cells were starved of IL-3 for 24 hours in a humidified incubator at 37°C in 5% CO_2 and air. Following IL-3 starvation the cells were plated out in 96 well culture dishes with or without diluted samples and cultured for 24 hrs in a cell culture incubator. 20 µl of serum free RPMI media containing 1 µCi of ^{3}H-thymidine was added to each well for the last 6-8 hours. The cells were then harvested on 96 well filter plates and washed with water. The filters were then counted.

Fig. 3 shows the effect of pronase, DTT and heat on the ability of APP to stimulate Ba/F3-*mpl* cell proliferation. For pronase digestion of APP, pronase (Boehringer Mannheim) or bovine serum albumin was coupled to Affi-gel10 (Biorad) and incubated individually with APP for 18hrs. at 37°C. Subsequently, the resins were

removed by centrifugation and supernatants assayed. APP was also heated to 80°C for 4 min. or made 100 µM DTT followed by dialysis against PBS.

Fig. 4 shows the elution of *mpl* ligand activity from Phenyl-Toyopearl, Blue-Sepharose and Ultralink-*mpl* columns. Fractions 4-8 from the *mpl* affinity column were the peak activity fractions eluted from the column.

Fig. 5 shows the SDS-PAGE of eluted Ultralink-*mpl* fractions. To 200 µl of each fraction 2-8, 1 ml of acetone containing 1mM HCl at -20°C was added. After 3hrs. at -20°C samples were centrifuged and resultant pellets were washed 2x with acetone at -20°C. The acetone pellets were subsequently dissolved in 30 µl of SDS-solubilization buffer, made 100 µM DTT and heated at 90°C for 5 min. The samples were then resolved on a 4-20% SDS-polyacrylamide gel and proteins were visualized by silver staining.

Fig. 6 shows elution of *mpl* ligand activity from SDS-PAGE. Fraction 6 from the *mpl*-affinity column was resolved on a 4-20% SDS-polyacrylamide gel under non-reducing conditions. Following electrophoresis the gel was sliced into 12 equal regions and electroeluted as described in the examples. The electroeluted samples were dialyzed into PBS and assayed at a 1/20 dilution. The Mr standards used to calibrate the gel were Novex Mark 12 standards.

Fig. 7 shows the effect of *mpl* ligand depleted APP on human megakaryocytopoiesis. *mpl* ligand depleted APP was made by passing 1 ml over a 1 ml *mpl*-affinity column (700 µg *mpl*-IgG/ml NHS-superose, Pharmacia). Human peripheral stem cell cultures were made 10% APP or 10% *mpl* ligand depleted APP and cultured for 12 days. Megakaryocytopoiesis was quantitated as described in the examples.

Fig. 8 shows the effect of *mpl*-IgG on the stimulation of human megakaryocytopoiesis by APP. Human peripheral stem cell cultures were made 10% with APP and cultured for 12 days. At day 0, 2 and 4, *mpl*-IgG (0.5 µg) or ANP-R-IgG (0.5 µg) was added. After 12 days megakaryocytopoiesis was quantitated as described in the examples. The average of duplicate samples is graphed with the actual duplicate data in parenthesis.

Fig. 9 shows both strands of a 390 bp fragment of human genomic DNA encoding the *mpl* ligand. The deduced amino acid sequence of "exon 3" (SEQ ID NO: 3), the coding sequence (SEQ ID NO: 4), and its compliment (SEQ ID NO: 5) are shown.

Fig. 10 shows deduced amino acid sequence of mature human *mpl* ligand (hML) (SEQ ID NO: 6) and mature human erythropoietin (hEPO) (SEQ ID NO: 7). The predicted amino acid sequence for the human *mpl* ligand is aligned with the human erythropoietin sequence. Identical amino acids are boxed and gaps introduced for optimal alignment are indicated by dashes. Potential N-glycosylation sites are underlined with a plain line for the hML and with a broken line for hEPO. The two cysteines important for erythropoietin activity are indicated by a large dot.

Fig. 11 shows deduced amino acid sequence of mature human *mpl* ligand isoforms hML (SEQ ID NO: 6), hML2 (SEQ ID NO: 8), hML3 (SEQ ID NO: 9), and hML4 (SEQ ID NO: 10). Identical amino acids are boxed and gaps introduced for optimal alignment are indicated by dashes.

Figs. 12A, 12B and **12C** show the effect of human *mpl* ligand on Ba/F3-*mpl* cell proliferation (**A**), *in vitro* human megakaryocytopoiesis quantitated using a radiolabeled murine IgG monoclonal antibody specific to the megakaryocyte glycoprotein $GPII_bIII_a$ (**B**), and murine thrombopoiesis measured in a platelet rebound assay (**C**).

Two hundred ninety-three cells were transfected by the $CaPO_4$ method (Gorman, C in *DNA Cloning : A New Approach* **2**:143-190 [1985]) with pRK5 vector alone, pRK5-hML or with pRK5-ML_{153} overnight (pRK5-ML_{153} was generated by introducing a stop codon after residue 153 of hML by PCR). Media was then conditioned for 36h and assayed for stimulation of cell proliferation of Ba/F3-*mpl* as described in **Example 1 (A)** or *in vitro* human megakaryocytopoiesis **(B)**. Megakaryocytopoiesis was quantitated using a ^{125}I radiolabeled murine IgG monoclonal antibody (HP1-1D) to the megakaryocyte specific glycoprotein $GPII_bIII_a$ as described (Grant *et al., Blood* **69**:1334-1339 [1987]). The effect of partially purified recombinant ML (rML) on *in vivo* platelet production (**C**) was determined using the rebound thrombocytosis assay described by McDonald, T.P. *Proc. Soc. Exp. Biol. Med.* **144**:1006-10012 (1973). Partially purified rML was prepared from 200ml of conditioned media containing the recombinant ML. The media was passed through a 2ml Blue-Separose column equilabrated in PBS and the column was washed with PBS and eluted with PBS containing 2M each of urea and NaCl. The active fraction was dialyzed into PBS and made 1mg/ml with endotoxin free BSA. The sample contained less than one unit of endotoxin /ml. Mice were injected with either 64,000, 32,000 or 16,000 units of rML or excipient alone. Each group consisted of six mice. The mean and standard deviation of each group is shown. p values were determined by a 2 tailed T-test comparing medians.

Fig. 13 compares the effect of human *mpl* ligand isoforms and variants in the Ba/F3-*mpl* cell proliferation assay. hML, mock, hML2, hML3, hML(R153A, R154A), and hML_{153} were assayed at various dilutions as described in **Example 1.**

Figs. 14A, 14B and **14C** show the deduced amino acid sequence (SEQ ID NO: 1) of human *mpl* ligand (hML) or human TPO (hTPO) and the human genomic DNA coding sequence (SEQ ID NO: 11). Nucleotides and amino acid residues are numbered at the beginning of each line.

Fig. 15 shows a SDS-PAGE of purified 293-$rhML_{332}$ and purified 293-$rhML_{153}$.

Fig. 16 shows the nucleotide sequence: cDNA coding (SEQ ID NO: 12) and deduced amino acid sequence (SEQ ID NO: 13) of the open reading frame of a murine ML isoform. This mature murine *mpl* ligand isoform contains 331 amino acid residues, four fewer than the putative full length mML, and is therefore designated mML2. Nucleotides are numbered at the beginning of each line. Amino acid residues are numbered above the sequence starting with Ser 1. The potential N-glycosylation sites are underlined. Cysteine residues are indicated by a dot above the sequence.

Fig. 17 shows the cDNA sequence (SEQ ID NO: 14) and predicted protein sequence (SEQ ID NO: 15) of this murine ML isoform (mML). Nucleotides are numbered at the beginning of each line. Amino acid residues are numbered above the sequence starting with Ser 1. This mature murine *mpl* ligand isoform contains 335 amino acid residues and is believed to be the full length *mpl* ligand, designated mML. The signal sequence is indicated with a dashed underline and the likely cleavage point is denoted with an arrow. The 5' and 3' untranslated regions are indicated with lower case letters. The two deleations found as a result of alternative splicing (mML2 and mML3) are underlined. The four cysteine residues are indicated by a dot. The seven potential N-glycosylation sites are boxed.

Fig. 18 compares the deduced amino acid sequence of the human ML isoform hML3 (SEQ ID NO: 9) and a murine ML isoform designated mML3 (SEQ ID NO: 16). The predicted amino acid sequence for the human *mpl* ligand is aligned with the murine *mpl* ligand sequence. Identical amino acids are boxed and gaps introduced for optimal alignment are indicated by dashes. Amino acids are numbered at the beginning of each line.

Fig. 19 compares the predicted amino acid sequences of mature ML isoforms from mouse-ML (SEQ ID NO: 17), porcine-ML (SEQ ID NO: 18) and human-ML (SEQ ID NO: 6). Amino acid sequences are aligned with gaps, indicated by dashes, introduced for optimal alingment. Amino acids are numbered at the beginning of each line with identical residues boxed. Potential N-glycosylation sites are indicated by a shaded box and cysteine residues are designated with a dot. The conserved di-basic amino acid motif that presents a potential protease cleavage site is underlined. The four amino acid deletion found to occur in all three species (ML2) is outlined with a bold box.

Fig. 20 shows the cDNA sequence (SEQ ID NO: 19) and predicted mature protein sequence (SEQ ID NO: 18) of a porcine ML isoform (pML). This porcine *mpl* ligand isoform contains 332 amino acid residues and is believed to be the full length porcine *mpl* ligand, designated pML. Nucleotides are numbered at the beginning of each line. Amino acid residues are numbered above the sequence starting with Ser 1.

Fig. 21 shows the cDNA sequence (SEQ ID NO: 20) and predicted mature protein sequence (SEQ ID NO: 21) of a porcine ML isoform (pML2). This porcine *mpl* ligand isoform contains 328 amino acid residues and is a four residues deletion form of the full length porcine *mpl* ligand, designated pML2. Nucleotides are numbered at the beginning of each line. Amino acid residues are numbered above the sequence starting with Ser 1.

Fig. 22 compares the deduced amino acid sequence of the full length porcine ML isoform pML (SEQ ID NO: 18) and a porcine ML isoform designated pML2 (SEQ ID NO: 21). The predicted amino acid sequence for the pML is aligned with pML2 sequence. Identical amino acids are boxed and gaps introduced for optimal alignment are indicated by dashes. Amino acids are numbered at the beginning of each line.

Fig. 23 shows the pertinent features of plasmid pSVI5.ID.LL.MLORF ("full length" or TPO_{332}) used to transfect host CHO-DP12 cells for production of CHO-rhTPO_{332}.

Fig. 24 shows the pertinent features of plasmid pSVI5.ID.LL.MLEPO-D ("truncated" or TPO_{153}) used to transfect host CHO-DP12 cells for production of CHO-rhTPO_{153}.

Figs. 25A, 25B, and **25C** show the effect of *E. coli*-rhTPO(Met^{-1}, 153) on platelets (**A**), red blood cells (**B**) and (**C**) white blood cells in normal mice. Two groups of 6 female C57 B6 mice were injected daily with either PBS buffer or 0.3µg *E. coli*-rhTPO(Met^{-1}, 153) (100µl sc.). On day 0 and on days 3-7 40µl of blood was

taken from the orbital sinus. This blood was immediately diluted in 10 ml of commercial diluant and complete blood counts were obtained on a Serrono Baker Hematology Analyzer 9018. The data are presented as means ± Standard error of the mean.

Figs. 26A, 26B and 26C show the effect of *E. coli*-rhTPO(Met^{-1}, 153) on platelets **(A)**, red blood cells **(B)** and **(C)** white blood cells in sublethally irradiated mice. Two groups of 10 female C57 B6 mice were sublethally irradiated with 750 cGy of gamma radiation from a ^{137}Cs source and injected daily with either PBS buffer or 3.0µg *E. coli*-rhTPO(Met^{-1}, 153) (100µl sc.). On day 0 and at subsequent intermediate time points 40µl of blood was taken from the orbital sinus. This blood was immediately diluted in 10 ml of commercial diluant and complete blood counts were obtained on a Serrono Baker Hematology Analyzer 9018. The data are presented as means ± Standard error of the mean.

Figs. 27A, 27B and 27C show the effect of CHO-rhTPO$_{332}$ on **(A)** platelets (thrombocytes), **(B)** red blood cells (erythrocytes) and **(C)** white blood cells (leukocytes) in normal mice. Two groups of 6 female C57 B6 mice were injected daily with either PBS buffer or 0.3µg CHO-rhTPO$_{332}$ (100µl sc.). On day 0 and on days 3-7 40µl of blood was taken from the orbital sinus. This blood was immediately diluted in 10 ml of commercial diluant and complete blood counts were obtained on a Serrono Baker Hematology Analyzer 9018. The data are presented as means ± Standard error of the mean.

Fig. 28 shows dose response curves for various forms of rhTPO obtained from various cell lines. Dose response curves were constructed to rhTPO from the following cell lines: hTPO$_{332}$ from CHO (full length from Chinese hamster ovary cells); hTPO$_{Met^{-1}\ 153}$ (*E. coli*-derived truncated form with an N-terminal methionine); hTPO$_{332}$ (full length TPO from human 293 cells); Met-less 155 *E-Coli* (the truncated form [rhTPO$_{155}$] without the terminal methionine from *E. coli*). Groups of 6 female C57B6 mice were injected daily for 7 days with rhTPO depending upon group. Each day 40µl of blood was taken from the orbital sinus for complete blood counts. The data presented above are the maximal effects seen with the various treatments and with the exception of (met 153 *E-Coli*) this occurred on day 7 of treatment. In the aforementioned "met 153 *E-Coli*" group the maximal effect was seen on day 5. The data are presented as means ± Standard error of the mean.

Fig. 29 shows dose response curves comparing the activity of full length and "cliped" forms of rhTPO produced in CHO cells with the truncated form from *E. coli*. Groups of 6 female C57B6 mice were injected daily with 0.3μg rhTPO of various types. On days 2-7 40μl of blood was taken from the orbital sinus for complete blood counts. Treatment groups were TPO_{153} the truncated form of TPO from *E. coli*; TPO_{332} (Mix fraction) Full length TPO containing approximately 80-90% full length and 10-20% clipped forms; TPO332(30K fraction) = purified clipped fraction from the original "mix" preparation; TPO332(70K fraction) = purified full length TPO fraction from the original "mix" preparation. The data are presented as means ± Standard error of the mean.

Fig. 30 is a cartoon showing the KIRA ELISA assay for measuring TPO. The figure shows the MPL/Rse.gD chimera and relavant parts of the parent receptors as well as the final construct (right portion of the figure) and a flow diagram (left portion of the figure) showing relavant steps of the assay.

Fig. 31 is a flow chart for the KIRA ELISA assay showing each step in the procedure.

Figs. 32A-32L provide the nucleotide sequence (SEQ ID NO: 22) of the pSVI17.ID.LL expression vector used for expression of Rse.gD in **Example 17.**

Fig. 33 is a schematic representation of the preperation of plasmid pMP1.

Fig. 34 is a schematic representation of the preperation of plasmid pMP21.

Fig. 35 is a schematic representation of the preperation of plasmid pMP151.

Fig. 36 is a schematic representation of the preperation of plasmid pMP202.

Fig. 37 is a schematic representation of the preperation of plasmid pMP172.

Fig. 38 is a schematic representation of the preperation of plasmid pMP210.

Fig. 39 is a table of the five best expressing TPO clones from the pMP210 plasmid bank (SEQ ID NOS: 23, 24, 25, 26, 27 and 28).

Fig. 40 is a schematic representation of the preperation of plasmid pMP41.

Fig. 41 is a schematic representation of the preperation of plasmid pMP57.

Fig. 42 is a schematic representation of the preperation of plasmid pMP251.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

In general, the following words or phrases have the indicated definition when used in the description, examples, and claims.

"Chaotropic agent" refers to a compound which, in aqueous solution and in suitable concentrations, can cause a change in the spatial configuration or conformation of a protein by at least partially disrupting the forces responsible for maintaining the normal secondary and tertiary structure of the protein. Such compounds include, for example, urea, guanidine·HCl, and sodium thiocyanate. High concentrations, usually 4-9M, of these compounds are normally required to exert the conformational effect on proteins.

"Cytokine" is a generic term for proteins released by one cell population which act on another cell as intercellular mediators. Examples of such cytokines are lymphokines, monokines, and traditional polypeptide hormones. Included among the cytokines are growth hormone, insulin-like growth factors, human growth hormone, N-methionyl human growth hormone, bovine growth hormone, parathyroid hormone, thyroxine, insulin, proinsulin, relaxin, prorelaxin, glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and leutinizing hormone (LH), hematopoietic growth factor, hepatic growth factor, fibroblast growth factor, prolactin, placental lactogen, tumor necrosis factor-α (TNF-α and TNF-β) mullerian-inhibiting substance, mouse gonadotropin-associated peptide, inhibin, activin, vascular endothelial growth factor, integrin, nerve growth factors such as NGF-β, platelet-growth factor, transforming growth factors (TGFs) such as TGF-α and TGF-β, insulin-like growth factor-I and -II, erythropoietin (EPO), osteoinductive factors, interferons such as interferon-α, -β, and -γ, colony stimulating factors (CSFs) such as macrophage-CSF (M-CSF), granulocyte-macrophage-CSF (GM-CSF), and granulocyte-CSF (G-CSF), interleukins (IL's) such as IL-1, IL-1α, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12 and other polypeptide factors including LIF, SCF, and kit-ligand. As used herein the foregoing terms are meant to include proteins from natural sources or from recombinant cell culture. Similarly, the terms are intended to include biologically active equivalents; *e.g.*, differing in amino acid sequence by one or more amino acids or in type or extent of glycosylation.

"*mpl* ligand", "*mpl* ligand polypeptide", "ML", "thrombopoietin" or "TPO" are used interchangeably herein and comprise any polypeptide that possesses the property of binding to *mpl*, a member of the cytokine receptor superfamily, and having a biological property of the ML as defined below. An exemplary biological property is the ability to stimulate the incorporation of labeled nucleotides (*e.g.*, ^{3}H-thymidine) into the DNA of IL-3 dependent Ba/F3 cells transfected with human *mpl* P. Another exemplary biological property is the ability to stimulate the incorporation of ^{35}S into circulating platelets in a mouse platelet rebound assay. This definition encompasses the polypeptide isolated from a *mpl* ligand source such as aplastic porcine plasma described herein or from another source, such as another animal species, including humans or prepared by recombinant or synthetic methods and includes variant forms including functional derivatives, fragments, alleles, isoforms and analogues thereof.

A "*mpl* ligand fragment" or "TPO fragment" is a portion of a naturally occurring mature full length *mpl* ligand or TPO sequence having one or more amino acid residues or carbohydrate units deleted. The deleted amino acid residue(s) may occur anywhere in the peptide including at either the N-terminal or C-terminal end or internally. The fragment will share at least one biological property in common with *mpl* ligand. *Mpl* ligand fragments typically will have a consecutive sequence of at least 10, 15, 20, 25, 30, or 40 amino acid residues that are identical to the sequences of the *mpl* ligand isolated from a mammal including the ligand isolated from aplastic porcine plasma or the human or murine ligand, especially the EPO-domain thereof. Representative examples of N-terminal fragments are hML_{153} or TPO(Met^{-1}1-153).

"*Mpl* ligand variants" or "*mpl* ligand sequence variants" as defined herein means a biologically active *mpl* ligand as defined below having less than 100% sequence identity with the *mpl* ligand isolated from recombinant cell culture or aplastic porcine plasma or the human ligand having the deduced sequence described in **Fig. 1** (SEQ ID NO: 1). Ordinarily, a biologically active *mpl* ligand variant will have an amino acid sequence having at least about 70% amino acid sequence identity with the *mpl* ligand isolated from aplastic porcine plasma or the mature murine or human ligand or fragments thereof (see **Fig. 1** [SEQ ID NO: 1]), preferably at least about 75%, more preferably at least about 80%, still more preferably at least about 85%, even more preferably at least about 90%, and most preferably at least about 95%.

A "chimeric *mpl* ligand" is a polypeptide comprising full length *mpl* ligand or one or more fragments thereof fused or bonded to a second heterologous polypeptide or one or more fragments thereof. The chimera will share at least one biological property in common with *mpl* ligand. The second polypeptide will typically be a cytokine, immunoglobin or fragment thereof.

"Isolated *mpl* ligand", "highly purified *mpl* ligand" and "substantially homogeneous *mpl* ligand" are used interchangeably and mean a *mpl* ligand that has been purified from a *mpl* ligand source or has been prepared by recombinant or synthetic methods and is sufficiently free of other peptides or proteins (1) to obtain at least 15 and preferably 20 amino acid residues of the N-terminal or of an internal amino acid sequence by using a spinning cup sequenator or the best commercially available amino acid sequenator marketed or as modified by published methods as of the filing date of this application, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Homogeneity here means less than about 5% contamination with other source proteins.

"Biological property" when used in conjunction with either the "*mpl* ligand" or "Isolated *mpl* ligand" means having thrombopoietic activity or having an *in vivo* effector or antigenic function or activity that is directly or indirectly caused or performed by a *mpl* ligand (whether in its native or denatured conformation) or a fragment thereof. Effector functions include*mpl* binding and any carrier binding activity, agonism or antagonism of *mpl*, especially transduction of a proliferative signal including replication, DNA regulatory function, modulation of the biological activity of other cytokines, receptor (especially cytokine) activation, deactivation, up- or down regulation, cell growth or differentiation and the like. An antigenic function means possession of an epitope or antigenic site that is capable of cross-reacting with antibodies raised against the native *mpl* ligand. The principal antigenic function of a *mpl* ligand polypeptide is that it binds with an affinity of at least about 10^6 l/mole to an antibody raised against the *mpl* ligand isolated from aplastic porcine plasma. Ordinarily, the polypeptide binds with an affinity of at least about 10^7 l/mole. Most preferably, the antigenically active *mpl* ligand polypeptide is a polypeptide that binds to an antibody raised against the *mpl* ligand having one of the above described effector functions. The antibodies used to define "biologically activity" are rabbit polyclonal antibodies raised by formulating the *mpl* ligand isolated from recombinant cell culture or aplastic porcine plasma in Freund's complete adjuvant, subcutaneously injecting the formulation, and boosting the immune response by intraperitoneal injection of the formulation until the titer of *mpl* ligand antibody plateaus.

"Biologically active" when used in conjunction with either the "*mpl* ligand" or "Isolated *mpl* ligand" means a *mpl* ligand or polypeptide that exhibits thrombopoietic activity or shares an effector function of the *mpl* ligand isolated from aplastic porcine plasma or expressed in recombinant cell culture described herein. A principal known effector function of the *mpl* ligand or polypeptide herein is binding to *mpl* and stimulating the incorporation of labeled nucleotides (^{3}H-thymidine) into the DNA of

IL-3 dependent Ba/F3 cells transfected with human *mpl* P. Another known effector function of the *mpl* ligand or polypeptide herein is the ability to stimulate the incorporation of ^{35}S into circulating platelets in a mouse platelet rebound assay. Yet another known effector function of *mpl* ligand is the ability to stimulate *in vitro* human megakaryocytopoiesis that may be quantitated by using a radio labeled monoclonal antibody specific to the megakaryocyte glycoprotein $GPIIb_bIII_a$.

"Percent amino acid sequence identity" with respect to the *mpl* ligand sequence is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the *mpl* ligand sequence isolated from aplastic porcine plasma or the murine or human ligand having the deduced amino acid sequence described in **Fig. 1** (SEQ ID NO: 1), after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. None of N-terminal, C-terminal, or internal extensions, deletions, or insertions into the *mpl* ligand sequence shall be construed as affecting sequence identity or homology. Thus exemplary biologically active *mpl* ligand polypeptides considered to have identical sequences include; prepro-*mpl* ligand, pro-*mpl* ligand, and mature *mpl* ligand.

"*Mpl* ligand microsequencing" may be accomplished by any appropriate standard procedure provided the procedure is sensitive enough. In one such method, highly purified polypeptide obtained from SDS gels or from a final HPLC step are sequenced directly by automated Edman (phenyl isothiocyanate) degradation using a model 470A Applied Biosystems gas phase sequencer equipped with a 120A phenylthiohydantion (PTH) amino acid analyzer. Additionally, *mpl* ligand fragments prepared by chemical (*e.g.*, CNBr, hydroxylamine, 2-nitro-5-thiocyanobenzoate) or enzymatic (*e.g.*, trypsin, clostripain, staphylococcal protease) digestion followed by fragment purification (*e.g.*, HPLC) may be similarly sequenced. PTH amino acids are analyzed using the ChromPerfect data system (Justice Innovations, Palo Alto, CA). Sequence interpretation is performed on a VAX 11/785 Digital Equipment Co. computer as described by Henzel *et al.*, *J. Chromatography,* **404**:41-52 [1987]. Optionally, aliquots of HPLC fractions may be electrophoresed on 5-20% SDS-PAGE, electrotransferred to a PVDF membrane (ProBlott, AIB, Foster City, CA) and stained with Coomassie Brilliant Blue (Matsurdiara, *J. Biol. Chem.,* **262**:10035-10038 [1987]. A specific protein identified by the stain is excised from the blot and N-terminal sequencing is carried out with the gas phase sequenator described above. For internal protein sequences, HPLC fractions are dried under vacuum (SpeedVac), resuspended in appropriate buffers, and digested with cyanogen bromide, the Lys-specific enzyme Lys-C (Wako Chemicals, Richmond, VA), or Asp-N (Boehringer Mannheim, Indianapolis, IN). After digestion, the resultant peptides are sequenced as a

mixture or after HPLC resolution on a C4 column developed with a propanol gradient in 0.1% TFA prior to gas phase sequencing.

"Thrombocytopenia" is defined as a platelet count below 150 X 10^9 per liter of blood.

"Thrombopoietic activity" is defined as biological activity that consists of accelerating the proliferation, differentiation and/or maturation of megakaryocytes or megakaryocyte precursors into the platelet producing form of these cells. This activity may be measured in various assays including an *in vivo* mouse platelet rebound synthesis assay, induction of platelet cell surface antigen assay as measured by an anti-platelet immunoassay (anti-$GPIIb_bIIIa$) for a human leukemia megakaryoblastic cell line (CMK), and induction of polyploidization in a megakaryoblastic cell line (DAMI).

"Thrombopoietin" (TPO) is defined as a compound having thrombopoietic activity or being capable of increasing serum platelet counts in a mammal. TPO is preferably capable of increasing endogenous platelet counts by at least 10%, more preferably by 50%, and most preferably capable of elevating platelet counts in a human to greater that 150X10^9 per liter of blood.

"Isolated *mpl* ligand nucleic acid" is RNA or DNA containing greater than 16 and preferably 20 or more sequential nucleotide bases that encode biologically active *mpl* ligand or a fragment thereof, is complementary to the RNA or DNA, or hybridizes to the RNA or DNA and remains stably bound under moderate to stringent conditions. This RNA or DNA is free from at least one contaminating source nucleic acid with which it is normally associated in the natural source and preferably substantially free of any other mammalian RNA or DNA. The phrase "free from at least one contaminating source nucleic acid with which it is normally associated" includes the case where the nucleic acid is present in the source or natural cell but is in a different chromosomal location or is otherwise flanked by nucleic acid sequences not normally found in the source cell. An example of isolated *mpl* ligand nucleic acid is RNA or DNA that encodes a biologically active *mpl* ligand sharing at least 75% sequence identity, more preferably at least 80%, still more preferably at least 85%, even more preferably 90%, and most preferably 95% sequence identity with the human, murine or porcine *mpl* ligand.

"Control sequences" when referring to expression means DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, a ribosome binding site, and possibly, other as yet poorly understood sequences. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

"Operably linked" when referring to nucleic acids means that the nucleic acids are placed in a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accord with conventional practice.

"Exogenous" when referring to an element means a nucleic acid sequence that is foreign to the cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is ordinarily not found.

"Cell," "cell line," and "cell culture" are used interchangeably herein and such designations include all progeny of a cell or cell line. Thus, for example, terms like "transformants" and "transformed cells" include the primary subject cell and cultures derived therefrom without regard for the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Mutant progeny that have the same function or biological activity as screened for in the originally transformed cell are included. Where distinct designations are intended, it will be clear from the context.

"Plasmids" are autonomously replicating circular DNA molecules possessing independent origins of replication and are designated herein by a lower case "p" preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from such available plasmids in accordance with published procedures. In addition, other equivalent plasmids are known in the art and will be apparent to the ordinary artisan.

"Restriction enzyme digestion" when referring to DNA means catalytic cleavage of internal phosphodiester bonds of DNA with an enzyme that acts only at certain locations or sites in the DNA sequence. Such enzymes are called "restriction endonucleases". Each restriction endonuclease recognizes a specific DNA sequence called a "restriction site" that exhibits two-fold symmetry. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors, and other requirements as established by the enzyme suppliers are used. Restriction enzymes commonly are designated by abbreviations composed of a capital

letter followed by other letters representing the microorganism from which each restriction enzyme originally was obtained and then a number designating the particular enzyme. In general, about 1 μg of plasmid or DNA fragment is used with about 1-2 units of enzyme in about 20 μl of buffer solution. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation of about 1 hour at 37°C is ordinarily used, but may vary in accordance with the supplier's instructions. After incubation, protein or polypeptide is removed by extraction with phenol and chloroform, and the digested nucleic acid is recovered from the aqueous fraction by precipitation with ethanol. Digestion with a restriction enzyme may be followed with bacterial alkaline phosphatase hydrolysis of the terminal 5' phosphates to prevent the two restriction-cleaved ends of a DNA fragment from "circularizing" or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Unless otherwise stated, digestion of plasmids is not followed by 5' terminal dephosphorylation. Procedures and reagents for dephosphorylation are conventional as described in sections 1.56-1.61 of Sambrook *et al., Molecular Cloning: A Laboratory Manual* [New York: Cold Spring Harbor Laboratory Press, 1989].

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally. For example, see Lawn *et al., Nucleic Acids Res.,* 9:6103-6114 [1981], and Goeddel *et al., Nucleic Acids Res.,* 8:4057 [1980].

"Southern analysis" or "Southern blotting" is a method by which the presence of DNA sequences in a restriction endonuclease digest of DNA or DNA-containing composition is confirmed by hybridization to a known, labeled oligonucleotide or DNA fragment. Southern analysis typically involves electrophoretic separation of DNA digests on agarose gels, denaturation of the DNA after electrophoretic separation, and transfer of the DNA to nitrocellulose, nylon, or another suitable membrane support for analysis with a radiolabeled, biotinylated, or enzyme-labeled probe as described in sections 9.37-9.52 of Sambrook *et al., supra.*

"Northern analysis" or "Northern blotting" is a method used to identify RNA sequences that hybridize to a known probe such as an oligonucleotide, DNA fragment, cDNA or fragment thereof, or RNA fragment. The probe is labeled with a radioisotope such as ^{32}P, or by biotinylation, or with an enzyme. The RNA to be analyzed is usually electrophoretically separated on an agarose or polyacrylamide gel, transferred to nitrocellulose, nylon, or other suitable membrane, and hybridized with the probe,

using standard techniques well known in the art such as those described in sections 7.39-7.52 of Sambrook *et al., supra.*

"Ligation" is the process of forming phosphodiester bonds between two nucleic acid fragments. For ligation of the two fragments, the ends of the fragments must be compatible with each other. In some cases, the ends will be directly compatible after endonuclease digestion. However, it may be necessary first to convert the staggered ends commonly produced after endonuclease digestion to blunt ends to make them compatible for ligation. For blunting the ends, the DNA is treated in a suitable buffer for at least 15 minutes at 15°C with about 10 units of the Klenow fragment of DNA polymerase I or T4 DNA polymerase in the presence of the four deoxyribonucleotide triphosphates. The DNA is then purified by phenol-chloroform extraction and ethanol precipitation. The DNA fragments that are to be ligated together are put in solution in about equimolar amounts. The solution will also contain ATP, ligase buffer, and a ligase such as T4 DNA ligase at about 10 units per 0.5 μg of DNA. If the DNA is to be ligated into a vector, the vector is first linearized by digestion with the appropriate restriction endonuclease(s). The linearized fragment is then treated with bacterial alkaline phosphatase or calf intestinal phosphatase to prevent self-ligation during the ligation step.

"Preparation" of DNA from cells means isolating the plasmid DNA from a culture of the host cells. Commonly used methods for DNA preparation are the large- and small-scale plasmid preparations described in sections 1.25-1.33 of Sambrook *et al., supra.* After preparation of the DNA, it can be purified by methods well known in the art such as that described in section 1.40 of Sambrook *et al., supra.*

"Oligonucleotides" are short-length, single- or double-stranded polydeoxynucleotides that are chemically synthesized by known methods (such as phosphotriester, phosphite, or phosphoramidite chemistry, using solid-phase techniques such as described in EP 266,032 published 4 May 1988, or via deoxynucleoside H-phosphonate intermediates as described by Froehler *et al., Nucl. Acids Res.,* **14**:5399-5407 [1986]). Further methods include the polymerase chain reaction defined below and other autoprimer methods and oligonucleotide syntheses on solid supports. All of these methods are described in Engels *et al., Agnew. Chem. Int. Ed. Engl.,* **28**:716-734 (1989). These methods are used if the entire nucleic acid sequence of the gene is known, or the sequence of the nucleic acid complementary to the coding strand is available. Alternatively, if the target amino acid sequence is known, one may infer potential nucleic acid sequences using known and preferred coding residues for each amino acid residue. The oligonucleotides are then purified on polyacrylamide gels.

"Polymerase chain reaction" or "PCR" refers to a procedure or technique in which minute amounts of a specific piece of nucleic acid, RNA and/or DNA, are amplified as described in U.S. Patent No. 4,683,195 issued 28 July 1987. Generally, sequence information from the ends of the region of interest or beyond needs to be available, such that oligonucleotide primers can be designed; these primers will be identical or similar in sequence to opposite strands of the template to be amplified. The 5' terminal nucleotides of the two primers may coincide with the ends of the amplified material. PCR can be used to amplify specific RNA sequences, specific DNA sequences from total genomic DNA, and cDNA transcribed from total cellular RNA, bacteriophage or plasmid sequences, *etc.* See generally Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, 51:263 [1987]; Erlich, ed., *PCR Technology*, (Stockton Press, NY, 1989). As used herein, PCR is considered to be one, but not the only, example of a nucleic acid polymerase reaction method for amplifying a nucleic acid test sample comprising the use of a known nucleic acid as a primer and a nucleic acid polymerase to amplify or generate a specific piece of nucleic acid.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% $NaDodSO_4$ (SDS) at 50°C, or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C. Another example is use of 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC and 0.1% SDS.

"Moderately stringent conditions" are described in Sambrook *et al.*, *supra*, and include the use of a washing solution and hybridization conditions (*e.g.*, temperature, ionic strength, and %SDS) less stringent than described above. An example of moderately stringent conditions are conditions such as overnight incubation at 37°C in a solution comprising: 20% formamide, 5 X SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 X Denhardt's solution, 10% dextran sulfate, and 20 µl/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 X SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength *etc.* as necessary to accommodate factors such as probe length and the like.

"Antibodies" (Abs) and "immunoglobulins" (Igs) are glycoproteins having the same structural characteristics. While antibodies exhibit binding specificity to a specific antigen, immunoglobulins include both antibodies and other antibody-like

molecules which lack antigen specificity. Polypeptides of the latter kind are, for example, produced at low levels by the lymph system and at increased levels by myelomas.

"Native antibodies and immunoglobulins" are usually heterotetrameric glycoproteins of about 150,000 daltons, composed of two identical light (L) chains and two identical heavy (H) chains. Each light chain is linked to a heavy chain by one covalent disulfide bond, while the number of disulfide linkages varies between the heavy chains of different immunoglobulin isotypes. Each heavy and light chain also has regularly spaced intrachain disulfide bridges. Each heavy chain has at one end a variable domain (V_H) followed by a number of constant domains. Each light chain has a variable domain at one and (V_L) and a constant domain at its other end; the constant domain of the light chain is aligned with the first constant domain of the heavy chain, and the light chain variable domain is aligned with the variable domain of the heavy chain. Particular amino acid residues are believed to form an interface between the light and heavy chain variable domains (Clothia *et al., J. Mol. Biol.*, **186**:651-663 [1985]; Novotny and Haber, *Proc. Natl. Acad. Sci. USA*, **82**:4592-4596 [1985]).

The term "variable" refers to the fact that certain portions of the variable domains differ extensively in sequence among antibodies and are used in the binding and specificity of each particular antibody for its particular antigen. However, the variability is not evenly distributed through the variable domains of antibodies. It is concentrated in three segments called complementarity determining regions (CDRs) or hypervariable regions both in the light chain and the heavy chain variable domains. The more highly conserved portions of variable domains are called the framework (FR). The variable domains of native heavy and light chains each comprise four FR regions, largely adopting a β-sheet configuration, connected by three CDRs, which form loops connecting, and in some cases forming part of, the β-sheet structure. The CDRs in each chain are held together in close proximity by the FR regions and, with the CDRs from the other chain, contribute to the formation of the antigen binding site of antibodies (see Kabat *et al., Sequences of Proteins of Immunological Interest*, National Institute of Health, Bethesda, MD [1987]). The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as participation of the antibody in antibody-dependent cellular toxicity.

Papain digestion of antibodies produces two identical antigen binding fragments, called "Fab" fragments, each with a single antigen binding site, and a residual "Fc" fragment, whose name reflects its ability to crystallize readily. Pepsin treatment yields an $F(ab')_2$ fragment that has two antigen combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen recognition and binding site. This region consists of a dimer of one heavy and one light chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen binding site on the surface of the V_H-V_L dimer. Collectively, the six CDRs confer antigen binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab' fragments differ from Fab fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. $F(ab')_2$ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other, chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda (λ), based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG and IgM, and several of these may be further divided into subclasses (isotypes), *e.g.*, IgG-1, IgG-2, IgG-3, and IgG-4; IgA-1 and IgA-2. The heavy chain constant domains that correspond to the different classes of immunoglobulins are called α, delta, epsilon, γ, and μ, respectively. The subunit structures and three-dimensional configurations of different classes of immunoglobulins are well known.

The term "antibody" is used in the broadest sense and specifically covers single monoclonal antibodies (including agonist and antagonist antibodies), antibody compositions with polyepitopic specificity, as well as antibody fragments (*e.g.*, Fab, $F(ab')_2$, and Fv), so long as they exhibit the desired biological activity.

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to conventional (polyclonal) antibody preparations which typically include different antibodies directed against different determinants (epitopes), each

monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they are synthesized by the hybridoma culture, uncontaminated by other immunoglobulins. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method first described by Kohler & Milstein, *Nature*, **256**:495 (1975), or may be made by recombinant DNA methods (see, *e.g.*, U.S. Patent No. 4,816,567 [Cabilly *et al.*]).

The monoclonal antibodies herein specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (U.S. Patent No. 4,816,567 (Cabilly *et al.*); and Morrison *et al.*, *Proc. Natl. Acad. Sci. USA*, **81**:6851-6855 [1984]).

"Humanized" forms of non-human (*e.g.*, murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibody may comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. These modifications are made to further refine and optimize antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details see: Jones

et al., *Nature*, 321:522-525 [1986]; Reichmann *et al.*, *Nature*, 332:323-329 [1988]; and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 [1992]).

"Non-immunogenic in a human" means that upon contacting the polypeptide in a pharmaceutically acceptable carrier and in a therapeutically effective amount with the appropriate tissue of a human, no state of sensitivity or resistance to the polypeptide is demonstratable upon the second administration of the polypeptide after an appropriate latent period (*e.g.*, 8 to 14 days).

II. Preferred Embodiments of the Invention

Preferred polypeptides of this invention are substantially homogeneous polypeptide(s), referred to as *mpl* ligand(s) or thrombopoietin (TPO), that possesse the property of binding to *mpl*, a member of the receptor cytokine superfamily, and having the biological property of stimulating the incorporation of labeled nucleotides (^{3}H-thymidine) into the DNA of IL-3 dependent Ba/F3 cells transfected with human *mpl* P. More preferred *mpl* ligand(s) are isolated mammalian protein(s) having hematopoietic, especially megakaryocytopoietic or thrombocytopoietic activity - namely, being capable of stimulating proliferation, maturation and/or differentiation of immature megakaryocytes or their predecessors into the mature platelet-producing form. Most preferred polypeptides of this invention are human *mpl* ligand(s) including fragments thereof having hematopoietic, megakaryocytopoietic or thrombopoietic activity. Optionally these human *mpl* ligand(s) lack glycosylation. Other prefered human *mpl* ligands are the "EPO-domain" of hML refered to as hML_{153} or $hTPO_{153}$, a truncated form of hML refered to as hML_{245} or $hTPO_{245}$ and the mature full length polypeptide having the amino acid sequence shown in **Fig. 1** (SEQ ID NO: 1), refered to as hML, hML_{332} or $hTPO_{332}$ and the biolocically active substitutional variant hML(R153A, R154A).

Optional preferred polypeptides of this invention are biologically or immunologically active *mpl* ligands variants selected from hML2, hML3, hML4, mML, mML2, mML3, pML and pML2.

Optional preferred polypeptides of this invention are biologically active *mpl* ligand variant(s) that have an amino acid sequence having at least 70% amino acid sequence identity with the human *mpl* ligand (see **Fig. 1** [SEQ ID NO: 1]), the murine *mpl* ligand (see **Fig. 16** [SEQ ID NOS: 12 & 13]), the recombinant porcine *mpl* ligand (see **Fig. 19** [SEQ ID NO: 18]) or the porcine *mpl* ligand isolated from aplastic porcine plasma, preferably at least 75%, more preferably at least 80%, still more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95%.

The *mpl* ligand isolated from aplastic porcine plasma has the following characteristics:

(1) The partially purified ligand elutes from a gel filtration column run in either PBS, PBS containing 0.1% SDS or PBS containing 4M $MgCl_2$ with Mr of 60,000-70,000;

(2) The ligand's activity is destroyed by pronase;

(3) The ligand is stable to low pH (2.5), SDS to 0.1%, and 2M urea;

(4) The ligand is a glycoprotein, based on its binding to a variety of lectin columns;

(5) The highly purified ligand elutes from non-reduced SDS-PAGE with a Mr of 25,000-35,000. Smaller amounts of activity also elute with Mr of ~18,000-22,000 and 60,000;

(6) The highly purified ligand resolves on reduced SDS-PAGE as a doublet with Mr of 28,000 and 31,000;

(7) The amino-terminal sequence of the 18,000-22,000, 28,000 and 31,000 bands is the same - SPAPPACDPRLLNKLLRDDHVLHGR (SEQ ID NO: 29); and

(8) The ligand binds and elutes from the following affinity columns

Blue-Sepharose,
CM Blue-Sepharose,
MONO-Q,
MONO-S,
Lentil lectin-Sepharose,
WGA-Sepharose,
Con A-Sepharose,
Ether 650m Toyopearl,
Butyl 650 m Toyopearl,
Phenyl 650m Toyopearl, and
Phenyl-Sepharose.

More preferred *mpl* ligand polypeptides are those encoded by human genomic or cDNA having an amino acid sequence described in **Fig. 1** (SEQ ID NO: 1).

Other preferred naturally occurring biologically active *mpl* ligand polypeptides of this invention include prepro-*mpl* ligand, pro-*mpl* ligand, mature *mpl* ligand, *mpl* ligand fragments and glycosylation variants thereof.

Still other preferred polypeptides of this invention include *mpl* ligand sequence variants and chimeras. Ordinarily, preferred *mpl* ligand sequence variants and chimeras are biologically active *mpl* ligand variants that have an amino acid sequence having at least 70% amino acid sequence identity with the human *mpl* ligand or the *mpl* ligand isolated from aplastic porcine plasma, preferably at least 75%, more

preferably at least 80%, still more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95%. An exemplary preferred *mpl* ligand variant is a N-terminal domain hML variant (refered to as the "EPO-domain" because of its sequence homology to erythropoietin). The prefered hML EPO-domain comprises about the first 153 amino acid residues of mature hML and is refered to as hML_{153}. An optionally prefered hML sequence variant comprises one in which one or more of the basic or dibasic amino acid residue(s) in the C-terminal domain is substituted with a non-basic amino acid residue(s) (*e.g.*, hydrophobic, neutral, acidic, aromatic, Gly, Pro and the like). A prefered hML C-terminal domain sequence variant comprises one in which Arg residues 153 and 154 are replaced with Ala residues. This variant is refered to as hML_{332}(R153A, R154A). An alternative prefered hML variant comprises either hML_{332} or hML_{153} in which amino residues 111-114 (QLPP or LPPQ) are deleted or replaced with a diferent tetrapeptide sequence(*e.g.* AGAG or the like). The foregoing deletion mutants are refered to as $\Delta 4hML_{332}$ or $\Delta 4hML_{153}$.

A preferred chimera is a fusion between *mpl* ligand or fragment (defined below) thereof with a heterologous polypeptide or fragment thereof. For example, hML_{153} may be fused to an IgG fragment to improve serum half-life or to IL-3, G-CSF or EPO to produce a molecule with inhanced thrombopoietic or chimeric hematopoietic activity.

An alternative preferred human *mpl* ligand chimera is a "ML-EPO domain chimera" that consists of the N-terminus 153 to 157 hML residues substituted with one or more, but not all, of the human EPO residues approximately aligned as shown in **Fig. 10** (SEQ ID NO: 7). In this embodiment, the hML chimera would be about 153-166 residues in length in which individual or blocks of residues from the human EPO sequence are added or substituted into the hML sequence at positions corresponding to the alignment shown in **Fig. 10** (SEQ ID NO: 6). Exemplary block sequence inserts into the N-terminus portion of hML would include one or more of the N-glycosylation sites at positions (EPO) 24-27, 38-40, and 83-85; one or more of the four predicted amphipathic α-helical bundles at positions (EPO) 9-22, 59-76, 90-107, and 132-152; and other highly conserved regions including the N-terminus and C-terminus regions and residue positions (epo) 44-52 (see *e.g.*, Wen *et al., Blood,* **82**:1507-1516 [1993] and Boissel *et al., J. Biol. Chem.,* **268(21)**:15983-15993 [1993]). It is contemplated this "ML-EPO domain chimera" will have mixed thrombopoietic-erythropoietic (TEPO) biological activity.

Other preferred polypeptides of this invention include *mpl* ligand fragments having a consecutive sequence of at least 10, 15, 20, 25, 30, or 40 amino acid residues that are identical to the sequences of the *mpl* ligand isolated from aplastic

porcine plasma or the human *mpl* ligand described herein (see *e.g.* **Table 14, Example 24**). A preferred *mpl* ligand fragment is human ML[1-X] where X is 153, 164, 191, 205, 207, 217, 229, or 245 (see **Fig. 1** [SEQ ID NO: 1] for the sequence of residues 1-X). Other preferred *mpl* ligand fragments include those produced as a result of chemical or enzymatic hydrolysis or digestion of the purified ligand.

Another preferred aspect of the invention is a method for purifying *mpl* ligand molecules comprises contacting a *mpl* ligand source containing the *mpl* ligand molecules with an immobilized receptor polypeptide, specifically *mpl* or a *mpl* fusion polypeptide, under conditions whereby the *mpl* ligand molecules to be purified are selectively adsorbed onto the immobilized receptor polypeptide, washing the immobilized support to remove non-adsorbed material, and eluting the molecules to be purified from the immobilized receptor polypeptide with an elution buffer. The source containing the *mpl* ligand may be plasma where the immobilized receptor is preferably a *mpl*-IgG fusion.

Alternatively, the source containing the *mpl* ligand is recombinant cell culture where the concentration of *mpl* ligand in either the culture medium or in cell lysates is generally higher than in plasma or other natural sources. In this case the above described *mpl*-IgG immunoaffinity method, while still useful, is usually not necessary and more traditional protein purification methods known in the art may be applied. Briefly, the preferred purification method to provide substantially homogeneous *mpl* ligand comprises: removing particulate debris, either host cells or lysed fragments by, for example, centrifugation or ultrafiltration; optionally, protien may be concentrated with a commercially available protein concentration filter; followed by separating the ligand from other impurities by one or more steps selected from; immunoaffinity, ion-exchange (*e.g.,* DEAE or matricies containing carboxymethyl or sulfopropyl groups), Blue-Sepharose, CM Blue-Sepharose, MONO-Q, MONO-S, lentil lectin-Sepharose, WGA-Sepharose, Con A-Sepharose, Ether Toypearl, Butyl Toypearl, Phenyl Toypearl, protein A Sepharose, SDS-PAGE, reverse phase HPLC (*e.g.,* silica gel with appended aliphatic groups) or Sephadex molecular seive or size exclusion chromatography, and ethanol or ammonium sulfate precipitation. A protease inhibitor such as methylsulfonylfluoride (PMSF) may be included in any of the foregoing steps to inhibit proteolysis.

In another preferred embodiment, this invention provides an isolated antibody capable of binding to the *mpl* ligand. A preferred *mpl* ligand isolated antibody is monoclonal (Kohler and Milstein, *Nature,* **256**:495-497 [1975]; Campbell, *Laboratory Techniques in Biochemistry and Molecular Biology*, Burdon *et al.*, Eds, Volume 13, Elsevier Science Publisrers, Amsterdam [1985]; and Huse *et al., Science,* **246**:1275-1281 [1989]). Preferred *mpl* ligand isolated antibody is one that binds

to *mpl* ligand with an affinity of at least about 10^6 l/mole. More preferably the antibody binds with an affinity of at least about 10^7 l/mole. Most preferably, the antibody is raised against the *mpl* ligand having one of the above described effector functions. The isolated antibody capable of binding to the *mpl* ligand may optionally be fused to a second polypeptide and the antibody or fusion thereof may be used to isolate and purify *mpl* ligand from a source as described above for immobilized *mpl* polypeptide. In a further preferred aspect of this embodiment, the invention provides a method for detecting the *mpl* ligand *in vitro* or *in vivo* comprising contacting the antibody with a sample, especially a serum sample, suspected of containing the ligand and detecting if binding has occurred.

In still further preferred embodiments, the invention provides an isolated nucleic acid molecule encoding the *mpl* ligand or fragments thereof, which nucleic acid molecule may be labeled or unlabeled with a detectable moiety, and a nucleic acid molecule having a sequence that is complementary to, or hybridizes under stringent or moderately stringent conditions with, a nucleic acid molecule having a sequence encoding a *mpl* ligand. A preferred *mpl* ligand nucleic acid is RNA or DNA that encodes a biologically active *mpl* ligand sharing at least 75% sequence identity, more preferably at least 80%, still more preferably at least 85%, even more preferably 90%, and most preferably 95% sequence identity with the human *mpl* ligand. More preferred isolated nucleic acid molecules are DNA sequences encoding biologically active *mpl* ligand, selected from: (a) DNA based on the coding region of a mammalian *mpl* ligand gene (*e.g.*, DNA comprising the nucleotide sequence provided in **Fig. 1** (SEQ ID NO: 2), or fragments thereof); (b) DNA capable of hybridizing to a DNA of (a) under at least moderately stringent conditions; and (c) DNA that is degenerate to a DNA defined in (a) or (b) which results from degeneracy of the genetic code. It is contemplated that the novel *mpl* ligands described herein may be members of a family of ligands or cytokines having suitable sequence identity that their DNA may hybridize with the DNA of **Fig. 1** (SEQ ID NO: 2) (or the complement or fragments thereof) under low to moderate stringency conditions. Thus a further aspect of this invention includes DNA that hybridizes under low to moderate stringency conditions with DNA encoding the *mpl* ligand polypeptides.

In a further preferred embodiment of this invention, the nucleic acid molecule is cDNA encoding the *mpl* ligand and further comprises a replicable vector in which the cDNA is operably linked to control sequences recognized by a host transformed with the vector. This aspect further includes host cells transformed with the vector and a method of using the cDNA to effect production of *mpl* ligand, comprising expressing the cDNA encoding the *mpl* ligand in a culture of the transformed host cells and recovering the *mpl* ligand from the host cell culture. The *mpl* ligand prepared in this manner is

preferably substantially homogeneous human *mpl* ligand. A preferred host cell for producing *mpl* ligand is Chinese hamster ovary (CHO) cells.

The invention further includes a preferred method for treating a mammal having an immunological or hematopoietic disorder, especially thrombocytopenia comprising administering a therapeutically effective amount of a *mpl* ligand to the mammal. Optionally, the *mpl* ligand is administered in combination with a cytokine, especially a colony stimulating factor or interleukin. Preferred colony stimulating factors or interleukins include; kit-ligand, LIF, G-CSF, GM-CSF, M-CSF, EPO, IL-1, IL-2, IL-3, IL-5, IL-6, IL-7, IL-8, IL-9 or IL-11.

III. Methods of Making

Platelet production has long been thought by some authors to be controlled by multiple lineage specific humoral factors. It has been postulated that two distinct cytokine activities, referred to as megakaryocyte colony-stimulating factor (meg-CSF) and thrombopoietin, regulate megakaryocytopoiesis and thrombopoiesis (Williams *et al., J. Cell Physiol.*, **110**:101-104 [1982]; Williams *et al., Blood Cells,* **15**:123-133 [1989]; and Gordon *et al., Blood,* **80**:302-307 [1992]). According to this hypothesis, meg-CSF stimulates the proliferation of progenitor megakaryocytes while thrombopoietin primarily affects maturation of more differentiated cells and ultimately platelet release. Since the 1960's the induction and appearance of both meg-CSF and thrombopoietin activities in the plasma, serum and urine of animals and humans following thrombocytopenic episodes has been well documented (Odell *et al., Proc. Soc. Exp. Biol. Med.*, **108**:428-431 [1961]; Nakeff *et al., Acta Haematol.*, **54**:340-344 [1975]; Specter, *Proc. Soc. Exp. Biol.*, **108**:146-149 [1961]; Schreiner *et al., J.Clin.Invest.*, **49**:1709-1713 [1970]; Ebbe, *Blood*, **44**:605-608 [1974]; Hoffman *et al., N. Engl. J. Med.*, **305**:533 [1981]; Straneva *et al., Exp. Hematol.*, **17**:1122-1127 [1988]; Mazur *et al., Exp. Hematol.*, **13**:1164 [1985]; Mazur *et al., J.Clin. Invest.*, **68**:733-741 [1981]; Sheiner *et al., Blood*, **56**:183-188 [1980]; Hill *et al., Exp. Hematol.*, **20**:354-360 [1992]; and Hegyi *et al., Int. J. Cell Cloning*, **8**:236-244 [1990]). These activities were reported to be lineage specific and distinct from known cytokines (Hill R.J. *et al., Blood* **80**:346 (1992); Erickson-Miller C.L. *et al., Brit. J. Haematol.*, **84**:197-203 (1993); Straneva J.E. *et al., Exp. Hematol.* **20**:4750(1992); and Tsukada J. *et al., Blood* **81**:866-867 [1993]). Heretofore, attempts to purify meg-CSF or thrombopoietin from thrombocytopenic plasma or urine have been unsuccessful.

Consistent with the above observations describing thrombocytopenic plasma, we have found that aplastic porcine plasma (APP) obtained from irradiated pigs stimulates human megakaryocytopoiesis *in vitro*. We have found that this stimulatory

activity is abrogated by the soluble extracellular domain of c-*mpl*, confirming APP as a potential source of the putative *mpl* ligand (ML). We have now successfully purified the *mpl* ligand from APP and amino acid sequence information was used to isolate murine, porcine and human ML cDNA. These ML's have sequence homology to erythropoietin and have both meg-CSF and thrombopoietin-like activities.

1. Purification and Identification of *mpl* Ligand from Plasma

As set forth above, aplastic plasma from a variety of species has been reported to contain activities that stimulate hematopoiesis *in vitro*, however no hematopoietic stimulatory factor has previously been reported isolated from plasma. One source of aplastic plasma is that obtained from irradiated pigs. This aplastic porcine plasma (APP) stimulates human hematopoiesis *in vitro*. To determine if APP contained the *mpl* ligand, its effect was assayed by measuring ^{3}H-thymidine incorporation into Ba/F3 cells transfected with human *mpl* P (Ba/F3-*mpl*) by the procedure shown in **Fig. 2**. APP stimulated ^{3}H-thymidine incorporation into Ba/F3-*mpl* cells but not Ba/F3 control cells (*i.e.*, not transfected with human *mpl* P). Additionally, no such activity was observed in normal porcine plasma. These results indicated that APP contained a factor or factors that transduced a proliferative signal through the *mpl* receptor and therefore might be the natural ligand for this receptor. This was futher supported by the finding that treatment of APP with soluble *mpl*-IgG blocked the stimulatory effects of APP on Ba/F3-*mpl* cells.

The activity in APP appeared to be a protein since pronase, DTT, or heat destroy the activity in APP (**Fig. 3**). The activity was also non-dialyzable. The activity was, however, stable to low pH (pH 2.5 for 2 hrs.) and was shown to bind and elute from several lectin-affinity columns, indicating that it was a glycoprotein. To further elucidate the structure and identity of this activity it was affinity purified from APP using a *mpl*-IgG chimera.

APP was treated according to the protocol set forth in **Examples 1 and 2**. Briefly, the *mpl* ligand was purified using hydrophobic interaction chromatography (HIC), immobilized dye chromatography, and *mpl*-affinity chromatography. The recovery of activity from each step is shown in **Fig. 4** and the fold purification is provided in **Table 1**. The overall recovery of activity through the *mpl*-affinity column was approximately 10%. The peak activity fraction (F6) from the *mpl*-affinity column has an estimated specific activity of 9.8 $x10^{6}$ units/mg. The overall purification from 5 liters of APP was approximately 4 $x10^{6}$ fold (0.8 units/mg to 3.3 x 10^{6} units/mg) with a 83 x 10^{6}-fold reduction in protein (250 gms to 3 μg). We estimated the specific activity of the ligand eluted from the *mpl*-affinity column to be ~$3x10^{6}$ units/mg.

TABLE 1

Purification of *mpl* Ligand

Sample	Volume mls	Protein mg/ml	Units/ml	Units	Specific Acitivity Units/mg	Yield %	Fold Purification
APP	5000	50	40	200,000	0.8	-	1
Phenyl	4700	0.8	40	200,000	50	94	62
Blue-Sep.	640	0.93	400	256,000	430	128	538
mpl (µl) (Fxns 5-7)	12	$5x10^{-4}$	1666	20,000	3,300,000	10	4,100,000

Protein was determined by the Bradford assay. Protein concentration of *mpl*-eluted fractions 5-7 are estimates based on staining intensity of a silver stained SDS-gel. One unit is defined as that causing 50% maximal stimulation of Ba/F3-*mpl* cell proliferation.

Analysis of eluted fractions from the *mpl* affinity column by SDS-PAGE (4-20%, Novex gel) run under reducing conditions, revealed the presence of several proteins (**Fig. 5**). Proteins that silver stained with the strongest intensity resolved with apparent Mr of 66,000, 55,000, 30,000, 28,000 and 18,000-22,000. To determine which of these proteins stimulated proliferation of Ba/F3-*mpl* cell cultures, the proteins were eluted from the gel as described in **Example 2.**

The results of this experiment showed that most of the activity eluted from a gel slice that included proteins with Mr 28,000-32,000, with lesser activity eluting in the 18,000-22,000 region of the gel (**Fig. 6**). The only proteins visible in these regions had Mr of 30,000, 28,000 and 18,000-22,000. To identify and obtain protein sequence for the proteins resolving in this region of the gel (*i.e.* bands at 30, 28 and 18-22 kDa), these three proteins were electroblotted to PVDF and sequenced as described in **Example 3.** Amino-terminus sequences obtained are provided in **Table 2.**

TABLE 2

Mpl Ligand Amino-Terminus Sequences

30 kDa 1 5 10 15 20 25 (S) P A P P A(C)D P R L L N K L L R D D (H/S) V L H (G) R L	(SEQ ID NO: 30)
28 kDa 1 5 10 15 20 25 (S) P A P P A X D P R L L N K L L R D D (H) V L (H) G R	(SEQ ID NO: 31)
18-22 kDa 1 5 10 X P A P P A X D P R L X (N) (K)	(SEQ ID NO: 32)

Computer-assisted analysis revealed these amino acid sequences to be novel. Because all three sequences were the same, it was believed the 30 kDa, 28 kDa and 18-22 kDa proteins were related and might be different forms of the same novel protein. Futhermore, this protein(s) was a likely candidate as the natural *mpl* ligand because the activity resolved on SDS-PAGE in the same region (28,000-32,000) of a 4-20% gel. In addition, the partially purified ligand migrated with a Mr of 17,000-30,000 when subjected to gel filtration chromatography using a Superose 12 (Pharmacia) column. It is believed the different Mr forms of the ligand are a result of proteolysis or glycosylation differences or other post or pre-translational modifications.

As described earlier, antisense human *mpl* RNA abrogated megakaryocytopolesis in human bone marrow cultures enriched with CD 34^+ progenitor cells without affecting the differentiation of other hematopoietic cell lineages (Methia *et al., supra*). This result suggested that the *mpl* receptor might play a role in the differentiation and proliferation of megakaryocytes *in vitro*. To further elucidate the role of the *mpl* ligand in megakaryocytopoiesis, the effects of APP and *mpl* ligand depleted APP on *in vitro* human megakaryocytopoiesis was compared. The effect of APP on human megakaryocytopoiesis was determined using a modification of the liquid suspension megakaryocytopoiesis assay described in **Example 4**. In this assay, human peripheral stem cells (PSC) were treated with APP before and after *mpl*-IgG affinity chromatography. $GPIIb_bIIIa$ stimulation of megakaryocytopoiesis was quantitated with an ^{125}I-anti-IIb_bIIIa antibody (**Fig. 7**). Shown in **Fig. 7**, 10% APP caused approximately a 3-fold stimulation while APP depleted of *mpl* ligand had no effect. Significantly, the *mpl* ligand depleted APP did not induce proliferation of the Ba/F3-*mpl* cells.

In another experiment, soluble human *mpl*-IgG added at days 0, 2 and 4 to cultures containing 10% APP neutralized the stimulatory effects of APP on human megakaryocytopoiesis (**Fig. 8**). These results indicate that the *mpl* ligand plays a role in regulating human megakaryocytopoiesis and therefore may be useful for the treatment of thrombocytopenia.

2. Molecular Cloning of the *mpl* Ligand

Based on the amino-terminal amino acid sequence obtained from the 30 kDa, 28 kDa and 18-22 kDa proteins (see **Table 2** above), two degenerate oligonucleotide primer pools were designed and used to amplify porcine genomic DNA by PCR. It was reasoned that if the amino-terminal amino acid sequence was encoded by a single exon then the correct PCR product was expected to be 69 bp long. A DNA fragment of this size was found and subcloned into pGEMT. The sequences of the oligonucleotide PCR primers and the three clones obtained are shown in **Example 5**. The amino acid sequence (PRLLNKLLR [SEQ ID NO: 33]) of the peptide encoded between the PCR primers was identical to that obtained by amino-terminal protein sequencing of the porcine ligand (see residues 9-17 for the 28 and 30 kDa porcine protein sequences above).

A synthetic oligonucleotide based on the sequence of the PCR fragment was used to screen a human genomic DNA library. A 45-mer oligonucleotide, designated pR45, was designed and synthesized based on the sequence of the PCR fragment. This oligonucleotide had the following sequence:

5' GCC-GTG-AAG-GAC-GTG-GTC-GTC-ACG-AAG-CAG-TTT-ATT-TAG-GAG-TCG 3'

(SEQ ID NO: 34)

This deoxyoligonucleotide was used to screen a human genomic DNA library in λgem12 under low stringency hybridization and wash conditions according to **Example 6.** Positive clones were picked, plaque purified and analyzed by restriction mapping and southern blotting. A 390 bp EcoRI-XbaI fragment that hybridized to the 45-mer was subcloned into pBluescript SK-. DNA sequencing of this clone confirmed that DNA encoding the human homolog of the porcine *mpl* ligand had been isolated. The human DNA sequence and deduced amino acid sequence are shown in **Fig. 9** (SEQ ID NOS: 3 & 4). The predicted positions of introns in the genomic sequence are also indicated by arrows, and define a putative exon ("exon 3").

Based on the human "exon 3" sequence (**Example 6**) oligonucleotides corresponding to the 3' and 5' ends of the exon sequence were synthesized. These 2 primers were used in PCR reactions employing as a template cDNA prepared from various human tissues. The expected size of the correct PCR product was 140 bp. After analysis of the PCR products on a 12% polyacrylamide gel, a DNA fragment of the

expected size was detected in cDNA libraries prepared from human adult kidney, 293 fetal kidney cells and cDNA prepared from human fetal liver.

A fetal liver cDNA library ($7x10^6$ clones) in lambda DR2 was next screened with the same 45-mer oligonucleotide used to screen the human genomic library and the fetal liver cDNA library under low stringency hybridization conditions. Positive clones were picked, plaque purified and the insert size was determined by PCR. One clone with a 1.8 kb insert was selected for further analysis. Using the procedures described in **Example 7** the nucleotide and deduced amino acid sequence of the human *mpl* ligand (hML) were obtained. These sequences are presented in **Fig. 1** (SEQ ID NOS: 1 & 2).

3. Structure of the Human *mpl* Ligand (hML)

The human *mpl* ligand (hML) cDNA sequence (**Fig. 1** [SEQ ID NO: 2]) comprises 1774 nucleotides followed by a poly(A) tail. It contains 215 nucleotides of 5' untranslated sequence and a 3' untranslated region of 498 nucleotides. The presumed initiation codon at nucleotide position (216-218) is within a consensus sequence favorable for eukaryotic translation initiation. The open reading frame is 1059 nucleotides long and encodes a 353 amino acid residue polypeptide, beginning at nucleotide position 220. The N-terminus of the predicted amino acid sequence is highly hydrophobic and probably corresponds to a signal peptide. Computer analysis of the predicted amino acid sequence (von Heijne *et al., Eur. J. Biochem.*, **133**:17-21 [1983]) indicates a potential cleavage site for signal peptidase between residues 21 and 22. Cleavage at that position would generate a mature polypeptide of 332 amino acid residues beginning with the amino-terminal sequence obtained from *mpl* ligand purified from porcine plasma. The predicted non-glycosylated molecular weight of the 332 amino acid residue ligand is about 38 kDa. There are 6 potential N-glycosylation sites and 4 cysteine residues.

Comparison of the *mpl* ligand sequence with the Genbank sequence database revealed 23% identity between the amino terminal 153 residues of mature human *mpl* ligand and human erythropoietin (**Fig. 10** [SEQ ID NOS: 6 & 7]). When conservative substitutions are taken into account, this region of hML shows 50% similarity to human erythropoietin (hEPO). Both hEPO and the hML contain four cysteines. Three of the 4 cysteines are conserved in hML, including the first and last cysteines. Site-directed mutagenesis experiments have shown that the first and last cysteines of erythropoietin form a disulfide bond that is required for function (Wang, F.F.*et al.*, *Endocrinology* **116**:2286-2292 [1983]). By analogy, the first and last cysteines of hML may also form a critical disulfide bond. None of the glycosylation sites are

conserved in hML. All potential hML N-linked glycosylation sites are located in the carboxy-terminal half of the hML polypeptide.

Similar to hEPO, the hML mRNA does not contain the consensus polyadenylation sequence AAUAAA, nor the regulatory element AUUUA that is present in 3' untranslated regions of many cytokines and is thought to influence mRNA stability (Shaw*et al., Cell,* **46**:659-667 [1986]). Northern blot analysis reveals low levels of a single 1.8 kb hML RNA transcript in both fetal and adult liver. After longer exposure, a weaker band of the same size could be detected in adult kidney. By comparison, human erythropoietin is expressed in fetal liver and, in response to hypoxia, the adult kidney and liver (Jacobs *et al., Nature,* **313**:804-809 [1985] and Bondurant *et al., Molec. Cell. Biol.,* **6**:2731-2733 [1986]).

The importance of the C-terminal region of the hML remains to be elucidated. Based on the presence of the six potential sites for N-linked glycosylation and the ability of the ligand to bind lectin-affinity columns, this region of the hML is likely glycosylated. In some gel elution experiments, we observed activity resolving with a M_r around 60,000 which may represent the full length, glycosylated molecule. The C-terminal region may therefore act to stabilize and increase the half-life of circulating hML. In the case of erythropoietin, the non-glycosylated form has full *in vitro* biological activity, but has a significantly reduced plasma half-life relative to glycosylated erythropoietin (Takeuchi *et al., J. Biol. Chem.,* **265**:12127-12130 [1990]; Narhi *et al., J. Biol. Chem.,* **266**:23022-23026 [1991] and Spivack *et al., Blood,* **7**:90-99 [1989]). The C-terminal domain of hML contains two di-basic amino acid sequences [Arg-Arg motifs at positions 153-154 and 245-246] that could serve as potential processing sites. Cleavage at these sites may be responsible for generating the 30, 28 and 18-22 kDa forms of the ML isolated from APP. Significantly, the Arg_{153}-Arg_{154} sequence occurs immediately following the erythropoietin-like domain of the ML. These observations indicate that full length ML may represent a precursor protein that undergoes limited proteolysis to generate the mature ligand.

4. Isoforms and Variants of the Human *mpl* Ligand

Isoforms or alternatively spliced forms of human *mpl* ligand were detected by PCR in human adult liver. Briefly, primers were synthesized corresponding to each end as well as selected internal regions of the coding sequence of hML. These primers were used in RT-PCR to amplify human adult liver RNA as described in **Example 10**. In addition to the full length form, designated hML, three other forms, designated hML2, hML3 and hML4, were observed or deduced. The mature deduced amino acid sequences of all four isoforms is presented in **Fig. 11** (SEQ ID NOS: 6, 8, 9 & 10).

hML3 has a 116 nucleotide deletion a position 700 which results in both an amino acid deletion and a frameshift. The cDNA now encodes a mature polypeptide that is 265 amino acid long and diverges from the hML sequence at amino acid residue 139. Finally, hML4 has both a 12 nucleotide deletion following nucleotide position 618 (also found in the mouse and the pig sequences [see below]) and the 116 bp deletion found in hML3. Altough no clones with only the 12 bp deletion (following nucleotide 619) have been isolated in the human (designated hML2), this form is likely to exist because such a isoform has been identified in both the mouse and pig (see below), and because it has been identified in conjunction with the116 nucleotide deletion in hML4.

Both a substitutional variant of hML in which the dibasic Arg_{153}-Arg_{154} sequence was replaced with two alanine residues and a "EPO-domain" truncated form of hML were constructed to determine whether the full length ML was necessary for biological activity. The Arg_{153}-Arg_{154} dibasic sequence substitutional variant, refered to as hML(R153A, R154A), was constructed using PCR as described in **Example 10.** The "EPO-domain" truncated form, hML_{153}, was also made using PCR by introducing a stop codon following Arg153.

5. Expression of Recombinant Human *mpl* Ligand (rhML) in Transiently Transfected Human Embryonic Kidney (293) Cells

To confirm that the cloned human cDNA encoded a ligand for *mpl*, the ligand was expressed in mammalian 293 cells under the control of the cytomegalovirus immediate early promoter using the expression vectors pRK5-hML or pRK5-hML_{153}. Supernatants from transiently transfected human embryonic kidney 293 cells were found to stimulate ^{3}H-thymidine incorporation in Ba/F3-*mpl* cells, but not in parental Ba/F3 cells (**Fig. 12A**). Media from the 293 cells transfected with the pRK vector alone did not contain this activity. Addition of *mpl*-IgG to the media abolished the stimulation (data not shown). These results show that the cloned cDNA encodes a functional human ML (hML).

To determine if the "EPO-domain" alone could bind and activate *mpl*, the truncated form of hML, $rhML_{153}$, was expressed in 293 cells. Supernatants from transfected cells were found to have activity similar to that present in supernatants from cells expressing the full length hML (**Fig. 12A**), indicating that the C-terminal domain of ML is not required for binding and activation of c-*mpl*.

6. *mpl* Ligand Stimulates Megakaryocytopoiesis and Thrombopoiesis

Both the full length rhML and the truncated $rhML_{153}$ forms of recombinant hML stimulated human megakaryocytopoiesis *in vitro* (**Fig. 12B**). This effect was observed in the absence of other exogenously added hematopoietic growth factors. With the exception of IL-3, the ML was the only hematopoietic growth factor tested that exhibited this activity. IL-11, IL-6, IL-1, erythropoietin, G-CSF, IL-9, LIF, kit ligand (KL), M-CSF, OSM and GM-CSF had no effect on megakaryocytopoiesis when tested separately in our assay (data not shown). This result demonstrates that the ML has megakaryocyte-stimulating activity, and indicates a role for ML in regulating megakaryocytopoiesis.

Thrombopoietic activities present in plasma of thrombocytopenic animals have been shown to stimulate platelet production in a mouse rebound thrombocytosis assay (McDonald, *Proc. Soc. Exp. Biol. Med.*, **14**:1006-1001 [1973] and McDonald *et al.*, *Scand. J. Haematol.*, **16**:326-334 [1976]). In this model mice are made acutely thrombocytopenic using specific antiplatelet serum, resulting in a predictable rebound thrombocytosis. Such immuno-thrombocythemic mice are more responsive to exogenous thrombopoietin-like activities than are normal mice (McDonald, *Proc. Soc. Exp. Biol. Med.*, **14**:1006-1001 [1973]), just as exhypoxic mice are more sensitive to erythropoietin than normal are mice (McDonald, *et al., J. Lab. Clin. Med.*, **77**:134-143 [1971]). To determine whether the rML stimulates platelet production *in vivo*, mice in rebound thrombocytosis were injected with partially purified rhML. Platelet counts and incorporation of ^{35}S into platelets were then quantitated. Injection of mice with 64,000 or 32,000 units of rML significantly increased platelet production, as evidenced by a ~20% increase in platelet counts ($p=0.0005$ and 0.0001, respectively) and a ~40% increase in ^{35}S incorporation into platelets ($p=0.003$) in the treated mice versus control mice injected with excipient alone (**Fig. 12C**). This level of stimulation is comparable to that which we have observed with IL-6 in this model (data not shown). Treatment with 16,000 units of rML did not significantly stimulate platelet production. These results indicate that ML stimulates platelet production in a dose-dependent manner and therefore possesses thrombopoietin-like activity.

293 cells were also transfected with the other hML isoform constructs described above and the supernatants were assayed using the Ba/F3-*mpl* proliferation assay (see **Fig. 13**). hML2 and hML3 showed no detectable activity in this assay, however the activity of hML(R153A, R154A) was similar to hML and hML_{153} indicating that processing at the Arg_{153}-Arg_{154} di-basic site is neither required for nor detrimental to activity.

7. Megakaryocytopoiesis and the *mpl* Ligand

It has been proposed that megakaryocytopoiesis is regulated at multiple cellular levels (Williams *et al.*, *J.Cell Physiol.*, **110**:101-104 [1982] and Williams *et al.*, *Blood Cells*, **15**:123-133 [1989]). This is based largely on the observation that certain hematopoietic growth factors stimulate proliferation of megakaryocyte progenitors while others appear to primarily affect maturation. The results presented here suggest that the ML acts both as a proliferative and maturation factor. That ML stimulates proliferation of megakaryocyte progenitors is supported by several lines of evidence. First, APP stimulates both proliferation and maturation of human megakaryocytes *in vitro*, and this stimulation is completely inhibited by *mpl*-IgG (**Figs. 7** and **8**). Furthermore, the inhibition of megakaryocyte colony formation by c-*mpl* antisense oligonucleotides (Methia *et al.*, *Blood*, **82**:1395-1401 [1993]) and the finding that c-*mpl* can transduce a proliferative signal in cells into which it is transfected (Skoda *et al.*, *EMBO*, **12**:2645-2653 [1993] and Vigon *et al.*, *Oncogene*, **8**:2607-2615 [1993]) also indicate that ML stimulates proliferation. The apparent expression of c-*mpl* during all stages of megakaryocyte differentiation (Methia *et al.*, *Blood*, **82**:1395-1401 [1993]) and the ability of recombinant ML to rapidly stimulate platelet production *in vivo* indicate that ML also affects maturation. The availability of recombinant ML makes possible a careful evaluation of its role in regulating megakaryocytopoiesis and thrombopoiesis as well as its potential to influence other hematopoietic lineages.

8. Isolation of the Human *mpl* Ligand (TPO) Gene

Human genomic DNA clones of the TPO gene were isolated by screening a human genomic library in λ-Gem12 with pR45, under low stringency conditions or under high stringency conditions with a fragment corresponding to the 3' half of human cDNA coding for the *mpl* ligand. Two overlapping lambda clones spanning 35 kb were isolated. Two overlapping fragments (BamH1 and EcoRI) containing the entire TPO gene were subcloned and sequenced (see **Figs. 14A, 14B and 14C**).

The structure of the human gene is composed of 6 exons within 7 kb of genomic DNA. The boundaries of all exon/intron junctions are consistent with the consensus motif established for mammalian genes (Shapiro, M. B., *et al.*, *Nucl. Acids Res.* **15**:7155 [1987]). Exon 1 and exon 2 contain 5' untranslated sequence and the initial four amino acids of the signal peptide. The remainder of the secretory signal and the first 26 amino acids of the mature protein are encoded within exon 3. The entire carboxyl domain and 3' untranslated as well as ~50 amino acids of the erythropoletin-

like domain are encoded within exon 6. The four amino acids involved in the deletion observed within hML-2 (hTPO-2) are encoded at the 5' end of exon 6.

Analysis of human genomic DNA by Southern blot indicated the gene for TPO is present in a single copy. The chromosomal location of the gene was determined by fluorescent *in situ* hybridization (FISH) which mapped to chromosome 3q27-28.

9. Expression and Purification of TPO from 293 Cells

Preperation and purification of ML or TPO from 293 cells is described in detail in **Example 19**. Briefly, cDNA corresponding to the TPO entire open reading frame was obtained by PCR using pRK5-h*mpl* I. The PCR product was purified and cloned between the restriction sites ClaI and XbaI of the plasmid pRK5tkneo (a pRK5 derived vector modified to express a neomycin resistance gene under the control of the thymidine kinase promote) to obtain the vector pRK5tkneo.ORF(a vector coding for the entire open reading frame).

A second vector coding for the EPO homologous domain was generated the same but using different PCR primers to obtain the final construct called pRK5-tkneoEPO-D.

These two constructs were transfected into Human Embryonic Kidney cells by the $CaPO_4$ method and neomycin resistant clones were selected and allowed to grow to confluency. Expression of ML_{153} or ML_{332} in the conditioned media from these clones was assessed using the Ba/F3-*mpl* proliferation assay.

Purification of $rhML_{332}$ was conducted as described in **Example 19**. Briefly, 293-$rhML_{332}$ conditioned media was applied to a Blue-Sepharose (Pharmacia) column that was subsequently washed with a buffer containing 2M urea. The column was eluted with a buffer containing 2M urea and 1M NaCl. The Blue-Sepharose elution pool was then directly applied to a WGA-Sepharose column, washed with 10 column volumes of buffer containing 2M urea and 1 M NaCl and eluted with the same buffer containing 0.5M N-acetyl-D-glucosamine. The WGA-Sepharose eluate was applied to a C4-HPLC column (Synchrom, Inc.) and eluted with a discontinuous propanol gradient. By SDS-PAGE the purified 293-$rhML_{332}$ migrates as a broad band in the 68-80 kDa region of the gel (see **Fig. 15**).

Purification of $rhML_{153}$ was also conducted as described in **Example 19**. Briefly, 293-$rhML_{153}$ conditioned media was resolved on Blue-Sepharose as described for $rhML_{332}$. The Blue Sepharose eluate was applied directly to a *mpl*-affinity column as described above. $RhML_{153}$ eluted from the *mpl*-affinity column was purified to homogeneity using a C4-HPLC column run under the same conditions used for $rhML_{332}$. By SDS-PAGE the purified $rhML_{153}$ resolves into 2 major and 2 minor bands with Mr of ~18,000-22,000 (see **Fig. 15**).

10. The Murine *mpl* Ligand

A DNA fragment corresponding to the coding region of the human *mpl* ligand was obtained by PCR, gel purified and labeled in the presence of ^{32}P-dATP and ^{32}P-dCTP. This probe was used to screen 10^6 clones of a mouse liver cDNA library in λGT10. A murine clone (**Fig. 16** [SEQ ID NOS: 12 & 13]) containing a 1443 base pair insert was isolated and sequenced. The presumed initiation codon at nucleotide position 138-141 was within a consensus sequence favorable for eukaryotic translation initiation (Kozak, M. *J.Cell Biol.*, **108**:229-241 [1989]). This sequence defines an open reading frame of 1056 nucleotides, which predicts a primary translation product of 352 amino acids. Flanking this open reading frame are 137 nucleotides of 5' and 247 nucleotides of 3' untranslated sequence. There is no poly(A) tail following the 3' untranslated region indicating that the clone is probably not complete. The N-terminus of the predicted amino acid sequence is highly hydrophobic and probably represents a signal peptide. Computer analysis (von Heijne, G. *Eur. J. Biochem.* **133**:17-21 [1983]) indicated a potential cleavage site for signal peptidase between residues 21 and 22. Cleavage at that position would generate a mature polypeptide of 331 amino acids (35 kDa) identified as mML_{331} (or mML2 for reasons described below). The sequence contains 4 cysteines, all conserved in the human sequence, and seven potential N-glycosylation sites, 5 of which are conserved in the human sequence. Again, as with hML, all seven potential N-glycosylation sites are located in the C-terminal half of the protein.

When compared with the human ML, considerable identity for both nucleotide and deduced amino acid sequences were observed in the "EPO-domains" of these ML's. However, when deduced amino acid sequences of human and mouse ML's were aligned, the mouse sequence appeared to have a tetrapeptide deletion between residues 111-114 corresponding to the 12 nucleotide deletion following nucleotide position 618 seen in both the human (see above) and pig (see below) cDNA's. Accordingly, additional clones were examined to detect possible murine ML isoforms. One clone encoded a 335 amino acid deduced sequence polypeptide containing the "missing" tetrapeptide LPLQ. This form is believed to be the full length murine ML and is refered to as mML or mML_{335}. The nucleotide and deduced amino acid sequence for mML are provided in **Fig. 17** (SEQ ID NOS: 14 & 15). This cDNA clone consists of 1443 base pairs followed by a poly(A) tail. It possesses an open reading frame of 1068 bp flanked by 134 bases of 5' and 241 bases of 3' untranslated sequence. The presumed initiation codon lies at nucleotide position 138-140. The open reading frame encodes a predicted protein of 356 amino acids, the first 21 of which are highly hydrophobic and likely function as a secretion signal.

Finally, a third murine clone was isolated, sequenced and was found to contained the 116 nucleotide deletion corresponding to hML3. This murine isoform is therefore denominated mML3. Comparison of the deduced amino acid sequences of these two isoforms is shown in **Fig. 18** (SEQ ID NOS: 9 & 16).

The overall amino acid sequence identity between human and mouse ML (**Fig. 19** [SEQ ID NOS: 6 & 17]) is 72% but this homology is not evenly distributed. The region defined as the "EPO-domain" (amino acids 1-153 for the human sequence and 1-149 for the mouse) is better conserved (86% homology) than the carboxy-terminal region of the protein (62% homology). This may further indicate that only the "EPO-domain" is important for the biological activity of the protein. Interestingly, of the two di-basic amino acid motifs found in hML, only the di-basic motif immediately following the "EPO-domain" (residue position 153-154) in the human sequence is present in the murine sequence. This is consistent with the possibility that the full length ML may represent a precursor protein that undergoes limited proteolysis to generate the mature ligand. Alternatively, proteolysis between Arg_{153}-Arg_{154} may facilitate hML clearance

An expression vector containing the entire coding sequence of mML was transiently transfected into 293 cells as described in **Example 1**. Conditioned media from these cells stimulated ^{3}H-thymidine incorporation into Ba/F3 cells expressing either murine or human *mpl* but had no effect on the parental (*mpl*-less) cell line. This indicates that the cloned murine ML cDNA encodes a functional ligand that is able to activate both the murine and human ML receptor (*mpl*).

11. The Porcine *mpl* Ligand

Porcine ML (pML) cDNA was isolated by RACE PCR as described in **Example 13**. A PCR cDNA product of 1342 bp was found in kidney and subcloned. Several clones were sequenced and found to encode a pig *mpl* ligand of 332 amino acid resudues referred to as pML (or pML_{332}) having the nucleotide and deduced amino acid sequence shown in **Fig. 20** (SEQ ID NOS: 18 & 19).

Again, a second form, designated pML2, encoding a protein with a 4 amino acid residue deletion (228 amino acid residues) was identified (see **Fig. 21** [SEQ ID NO: 21]). Comparison of pML and pML2 amino acid sequences shows the latter form is identical except that the tetrapeptide QLPP corresponding to residues 111-114 inclusive have been deleted (see **Fig. 22** [SEQ ID NOS: 18 & 21]). The four amino acid deletions observed in both murine and porcine ML cDNA occur at precisely the same position within the predicted proteins.

Comparison of the predicted amino acid sequences of the mature ML from human, mouse, and pig (**Fig. 19** [SEQ ID NOS: 6, 17 & 18]) indicates that overall

sequence identity is 72 percent between mouse and human, 68 percent between mouse and pig and 73 percent between pig and human. The homology is substantially greater in the amino-terminal half of the ML (EPO homologous domain). This domain is 80 to 84 percent identical between any two species whereas the carboxy-terminal half (carbohydrate domain) is only 57 to 67 percent identical. A di-basic amino acid motif that could represent a protease cleavage site is present at the carboxyl end of the erythropoeitin homology domain. This motif is conserved between the three species at this position (**Fig. 19** [SEQ ID NOS: 6, 17 & 18]). A second di-basic site present at position 245 and 246 in the human sequence is not present in the mouse or pig sequences. The murine and the pig ML sequence contain 4 cysteines, all conserved in the human sequence. There are seven potential N-glycosylation sites within the mouse ligand and six within the porcine ML, 5 of which are conserved within the human sequence. Again, all the potential N-glycosylation sites are located in the C-terminal half of the protein.

12. Expression and Purification of TPO from Chinese Hamster Ovary (CHO) Cells

The expression vectors used to transfect CHO cells are designated: pSVI5.ID.LL.MLORF (full length or TPO_{332}), and pSVI5.ID.LL.MLEPO-D (truncated or TPO_{153}). The pertinent features of these plasmids are presented in **Fig. 23** and **24**.

The transfection procedures are described in **Example 20**. Briefly, cDNA corresponding to the entire open reading frame of TPO was obtained by PCR. The PCR product was purified and cloned between two restriction sites (ClaI and SalI) of the plasmid pSVI5.ID.LL to obtain the vector pSVI5.ID.LL.MLORF. A second construct corresponding to the EPO homologous domain was generated the same way but using a different reverse primer(EPOD.Sal). The final construct for the vector coding for the EPO homologous domain of TPO is called pSVI5.ID.LL.MLEPO-D.

These two constructs were linearized with NotI and transfected into Chinese Hamster Ovary Cells (CHO-DP12 cells, EP 307,247 published 15 March 1989) by electroporation. 10^7 cells were electroporated in a BRL electroporation apparatus (350 Volts, 330 mF, low capacitance) in the presence of 10, 25 or 50 mg of DNA as described (Andreason, G.L. *J. Tissue Cult. Meth.* 15,56 [1993]). The day following transfection, cells were split in DHFR selective media (High glucose DMEM-F12 50:50 without glycine, 2mM glutamine, 2-5% dialyzed fetal calf serum). 10 to 15 days later individual colonies were transferred to 96 well plates and allowed to grow to confluency. Expression of ML_{153} or ML_{332} in the conditioned media from these clones was assessed using the Ba/F3-*mpl* proliferation assay (described in **Example 1**).

The process for purifying and isolating TPO from harvested CHO cell culture fluid is described in **Example 20**. Briefly, harvested cell culture fluid (HCCF) is applied to a Blue Sepharose column (Phamacia) at a ratio of approximately 100L of HCCF per liter of resin. The column is then washed with 3 to 5 column volumes of buffer followed by 3 to 5 column volumes of a buffer containing 2.0M urea. TPO is then eluted with 3 to 5 column volumes of buffer containing both 2.0M urea and 1.0M NaCl.

The Blue Sepharose eluate pool containing TPO is then applied to a Wheat Germ Lectin Sepharose column (Pharmacia) equilibrated in the Blue Sepharose eluting buffer at a ratio of from 8 to 16 ml of Blue Sepharose eluate per ml of resin. The column is then washed with 2 to 3 column volumes of equilibration buffer. TPO is then eluted with 2 to 5 column volumes of a buffer containing 2.0M urea and 0.5M N-acetyl-D-glucosamine.

The Wheat Germ Lectin eluate containing TPO is then acidified and $C_{12}E_8$ is added to a final concentration of 0.04%. The resulting pool is applied to a C4 reversed phase column equilibrated in 0.1% TFA, 0.04% $C_{12}E_8$ at a load of approximately 0.2 to 0.5 mg protein per ml of resin.

The protein is eluted in a two phase linear gradient of acetonitrile containing 0.1% TFA and 0.04% $C_{12}E_8$ and a pool is made on the basis of SDS-PAGE.

The C4 Pool is then diluted and diafilitered versus approximately 6 volumes of buffer on an Amicon YM or like ultrafiltration membrane having a 10,000 to 30,000 Dalton molecular weight cut-off. The resulting diafiltrate may be then directly processed or further concentrated by ultrafiltration. The diafiltrate/concentrate is usually adjusted to a final concentration of 0.01% Tween-80.

All or a portion of the diafiltrate/concentrate equivalent to 2 to 5% of the calculated column volume is then applied to a Sephacryl S-300 HR column (Pharmacia) equilibrated in a buffer containing 0.01% Tween-80 and chromatographed. The TPO containing fractions which are free of aggregate and proteolytic degradation products are then pooled on the basis of SDS-PAGE. The resulting pool is filtered and stored at 2-8°C.

13. Methods for Transforming and Inducing TPO Synthesis in a Microorganism and Isolating, Purifying and Refolding TPO Made Therein

Construction of *E. coli* TPO expression vectors is described in detail in **Example 21**. Briefly. plasmids pMP21, pMP151, pMP41, pMP57 and pMP202 were all designed to express the first 155 amino acids of TPO downstream of a small leader which varies among the different constructs. The leaders provide primarily for

high level translation initiation and rapid purification. The plasmids pMP210-1, -T8, -21, -22, -24, -25 are designed to express the first 153 amino acids of TPO downstream of an initiation methionine and differ only in the codon usage for the first 6 amino acids of TPO, while the plasmid pMP251 is a derivative of pMP210-1 in which the carboxy-terminal end of TPO is extended by two amino acids. All of the above plasmids will produce high levels of intracellular expression of TPO in *E. coli* upon induction of the tryptophan promoter (Yansura, D. G. *et. al. Methods in Enzymology* (Goeddel, D. V., Ed.) **185**:54-60, Academic Press, San Diego [1990]). The plasmids pMP1 and pMP172 are intermediates in the construction of the above TPO intracellular expression plasmids.

The above TPO expression plasmids were used to transform the *E. coli* using the $CaCl_2$ heat shock method (Mandel, M. *et al. J. Mol. Biol.*, **53**:159-162, [1970]) and other procedures described in **Example 21**. Briefly, the transformed cells were grown first at 37°C until the optical density (600nm) of the culture reached approximately 2-3. The culture was then diluted and, after growth with aeration, acid was added. The culture was then allowed to continue growing with aeration for another 15 hours after which time the cells were harvested by centrifugation.

The Isolation, Purification and Refolding procedures given below for production of biologically active, refolded human TPO or fragments thereof is described in **Examples 22** and **23** can be applied for the recovery of any TPO variant including N and C terminal extended forms. Other procedures suitable for refolding recombinant or synthetic TPO can be found in the following patents; Builder *et al.*, U.S. Patent 4,511,502; Jones *et al.*, U.S. Patent 4,512,922; Olson U.S. Patent 4,518,526 and Builder *et al.*, U.S. Patent 4,620,948; for a general description of the recovery and refolding process for a variety of recombinant proteins expressed in an insoluble form in *E. coli.*

A. Recovery of non-soluble TPO

A microorganism such as *E. coli* expressing TPO encoded by any suitable plasmid is fermented under conditions in which TPO is deposited in insoluble "refractile bodies". Optionally, cells are first washed in a cell disruption buffer. Typically, about 100g of cells are resuspended in about 10 volumes of a cell disruption buffer (*e.g.* 10 mM Tris, 5 mM EDTA, pH 8) with, for example, a Polytron homogenizer and the cells centrifuged at 5000 x g for 30 minutes. Cells are then lysed using any conventional technique such as tonic shock, sonication, pressure cycling, chemical or enzymatic methods. For example, the washed cell pellet above may be resuspended in another 10 volumes of a cell disruption buffer with a homogenizer and the cell suspension is passed through an LH Cell Disrupter (LH Inceltech, Inc.) or through a Microfluidizer (Microfluidics International) according to the manufactures'

instructions. The particulate matter containing TPO is then separated from the liquid phase and optionally washed with any suitable liquid. For example, a suspension of cell lysate may be centrifuged at 5,000 X g for 30 minutes, resuspended and optionally centrifuged a second time to make a washed refractile body pellet. The washed pellet may be used immediately or optionally stored frozen (at *e.g.* -70°C).

B. Solubilization and Purification of Monomeric TPO

Insoluble TPO in the refractile body pellet is then solubilized with a solublizing buffer. The solublizing buffer contains a chaotropic agent and is usually buffered at a basic pH and contains a reducing agent to improve the yield of monomeric TPO. Representative chaotropic agents include urea, guanidine·HCl, and sodium thiocyanate. A preferred chaotropic agent is guanidine·HCl. The concentration of chaotropic agent is usually 4-9M, preferably 6-8M. The pH of the solublizing buffer is maintained by any suitable buffer in a pH range of from about 7.5-9.5, preferably 8.0-9.0 and most preferably 8.0. Preferably the solubilizing buffer also contains a reducing agent to aid formation of the monomeric form of TPO. Suitable reducing agents include organic compounds containing a free thiol (RSH). Representative reducing agents include dithiothreitol (DTT), dithioerythritol (DTE), mercaptoethanol, glutathione (GSH), cysteamine and cysteine. A preferred reducing agent is dithiothreitol (DTT). Optionally, the solubilizing buffer may contain a mild oxidizing agent (*e.g.* molecular oxygen) and a sulfite salt to form monomeric TPO via sulfitolysis. In this embodiment, the resulting TPO-S-sulfonate is later refolded in the presence of the redox buffer (*e.g.* GSH/GSSG) to form the properly folded TPO.

The TPO protein is usually further purified using, for example, centrifugation, gel filtration chromatography and reversed phase column chromatography.

By way of illustration, the following procedure has produced suitable yields of monomeric TPO. The refractile body pellet is resuspended in about 5 volumes by weight of the solubilizing buffer (20 mM Tris, pH 8, with 6-8 M guanidine and 25 mM DTT) and stirred for 1-3 hr., or overnight, at 4°C to effect solubilization of the TPO protein. High concentrations of urea (6-8M) are also useful but generally result in somewhat lower yields compared to guanidine. After solubilization, the solution is centrifuged at 30,000 x g for 30 min. to produce a clear supernatant containing denatured, monomeric TPO protein. The supernatant is then chromatographed on a Superdex 200 gel filtration column (Pharmacia, 2.6 x 60 cm) at a flow rate of 2 ml/min. and the protein eluted with 20 mM Na phosphate, pH 6.0, with 10 mM DTT. Fractions containing monomeric, denatured TPO protein eluting between 160 and 200 ml are pooled. The TPO protein is further purified on a semi-preparative C4 reversed phase column (2 x 20 cm VYDAC). The sample is applied at 5 ml/min. to a column equilibrated in 0.1% TFA (trifluoroacetic acid) with 30% acetonitrile. The protein is

eluted with a linear gradient of acetonitrile (30-60% in 60 min.). The purified reduced protein elutes at approximately 50% acetonitrile. This material is used for refolding to obtain biologically active TPO variant.

C. Refolding TPO to Generate the Biologically Active Form

Following solubilization and further purification of TPO, the biologically active form is obtained by refolding the denatured monomeric TPO in a redox buffer. Because of the high potency of TPO (half maximal stimulation in the Ba/F3 assay is achieved at approximately 3 pg/ml), it is possible to obtain biologically active material utilizing many different buffer, detergent and redox conditions. However, under most conditions only a small amount of properly folded material (<10%) is obtained. For commercial manufacturing processes, it is desirable to have refolding yields at least 10%, more preferably 30-50% and most preferably >50%. Many different detergents including Triton X-100, dodecyl-beta-maltoside, CHAPS, CHAPSO, SDS, sarkosyl, Tween 20 and Tween 80, Zwittergent 3-14 and others were found suitable for producing at least some properly folded material. Of these however, the most preferred detergents were those of the CHAPS family (CHAPS and CHAPSO) which were found to work best in the refolding reaction and to limit protein aggregation and improper disulfide formation. Levels of CHAPS greater than about 1% were most preferred. Sodium chloride was required for the best yields, with the optimal levels between 0.1 M and 0.5M. The presence of EDTA (1-5 mM) in the redox buffer was preferred to limit the amount of metal-catalyzed oxidation (and aggregation) which was observed with some preparations. Glycerol concentrations of greater than 15% produced the optimal refolding conditions. For maximum yields, it was essential to have a redox pair in the redox buffer consisting of both an oxidized and reduced organic thiol (RSH). Suitable redox pairs include mercaptoethanol, glutathione (GSH), cysteamine, cysteine and their corresponding oxidized forms. Preferred redox pairs were glutathione(GSH):oxidized glutathione(GSSG) or cysteine:cystine. The most preferred redox pair was glutathione(GSH):oxidized glutathione(GSSG). Generally higher yields were observed when the mole ratio of oxidized member of the redox pair was equal to or in excess over the reduced member of the redox pair. pH values between 7.5 and about 9 were optimal for refolding of these TPO variants. Organic solvents (*e.g.* ethanol, acetonitrile, methanol) were tolerated at concentrations of 10-15% or lower. Higher levels of organic solvents increased the amount of improperly folded forms. Tris and phosphate buffers were generally useful. Incubation at 4 °C also produced higher levels of properly folded TPO.

Refolding yields of 40-60% (based on the amount of reduced and denatured TPO used in the refolding reaction) are typical for preparations of TPO that have been purified through the first C4 step. Active material can be obtained when less pure

preparations (*e.g.* directly after the Superdex 200 column or after the initial refractile body extraction) although the yields are less due to extensive precipitation and interference of non-TPO proteins during the TPO refolding process.

Since TPO contains 4 cysteine residues, it is possible to generate three different disulfide versions of this protein:

version 1: disulfides between cysteine residues 1-4 and 2-3
version 2: disulfides between cysteine residues 1-2 and 3-4
version 3: disulfides between cysteine residues 1-3 and 2-4.

During the initial exploration in determining refolding conditions, several different peaks containing the TPO protein were separated by C4 reversed phase chromatography. Only one of these peaks had significant biological activity as determined using the Ba/F3 assay. Subsequently, the refolding conditions were optimized to yield preferentially that version. Under these conditions, the misfolded versions were less than 10-20% of the total monomeric TPO obtained from the solubilizing step.

The disulfide pattern for the biologically active TPO has been determined to be 1-4 and 2-3 by mass spectrometry and protein sequencing, where the cysteines are numbered sequentially from the amino-terminus. This cysteine cross-linking pattern is consistent with the known disulfide bonding pattern of the related molecule erythropoietin.

D. Biological Activity of Recombinant, Refolded TPO

Refolded and purified TPO has activity in both *in vitro* and *in vivo* assays. For example, in the Ba/F3 assay, half-maximal stimulation of thymidine incorporation into the Ba/F3 cells for TPO (Met^{-1} 1-153) was achieved at 3.3 pg /ml (0.3 pM). In the *mpl* receptor-based ELISA, half-maximal activity occurred at 1.9 ng/ml (120 pM). In normal and myelosuppressed animals produced by near-lethal X-radiation, refolded TPO (Met^{-1} 1-153) was highly potent (activity was seen at doses as low as 30 ng/mouse) to stimulate the production of new platelets. Similar biological activity was observed for other forms of TPO refolded in accordance with the above described procedures (see **Figs. 25, 26 and 28**).

14. Methods for Measurement of Thrombopoietic Activity

Thrombopoietic activity may be measured in various assays including the Ba/F3 *mpl* ligand assay described in **Example 1**, an *in vivo* mouse platelet rebound synthesis assay, induction of platelet cell surface antigen assay as measured by an anti-platelet immunoassay (anti-$GPIIb_bIIIa$) for a human leukemia megakaryoblastic cell line (CMK) (see Sato *et al.*, *Brit. J. Heamatol.*, **72**:184-190 [1989])(see also the liquid suspension megakaryocytopoiesis assay described in **Example 4**), and

induction of polyploidization in a megakaryoblastic cell line (DAMI) (see Ogura *et al.*, *Blood*, **72(1)**:49-60 [1988]). Maturation of megakaryocytes from immature, largely non-DNA synthesizing cells, to morphologically identifiable megakaryocytes involves a process that includes appearance of cytoplasmic organelles, acquisition of membrane antigens ($GPIIb_bIIIa$), endoreplication and release of platelets as described in the background. A lineage specific promoter (*i.e.*, the *mpl* ligand) of megakaryocyte maturation would be expected to induce at least some of these changes in immature megakaryocytes leading to platelet release and alleviation of thrombocytopenia. Thus, assays were designed to measure the emergence of these parameters in immature megakaryocyte cell lines, *i.e.*, CMK and DAMI cells. The CMK assay (**Example 4**) measures the appearance of a specific platelet marker, $GPIIb_bIIIa$, and platelet shedding. The DAMI assay (**Example 15**) measures endoreplication since increases in ploidy are hallmarks of mature megakaryocytes. Recognizable megakaryocytes have ploidy values of 2N, 4N, 8N, 16N, 32N, *etc.* Finally, the *in vivo* mouse platelet rebound assay (**Example 16**) is useful in demonstrating that administration of the test compound (here the *mpl* ligand) results in elevation of platelet numbers.

Two additional *in vitro* assays have been developed to measure TPO activity. The first is a kinase receptor activation (KIRA) ELISA in which CHO cells are transfected with a *mpl*-Rse chimera and tyrosine phosphorylation of Rse is measured by ELISA after exposure of the *mpl* portion of the chimera to *mpl* ligand (see **Example 17**). The second is a receptor based ELISA in which ELISA plate coated rabbit anti-human IgG captures human chimeric receptor *mpl*-IgG which binds the *mpl* ligand being assayed. A biotinylated rabbit polyclonal antibody to *mpl* ligand (TPO_{155}) is used to detect bound *mpl* ligand which is measured using streptavidin-peroxidase as described in **Example 18**.

15. *In Vivo* Biological Response of Normal and Sublethally Irradiated Mice Treated with TPO

Both normal and sublethally irradiated mice were treated with truncated and full length TPO isolated from Chinese hamster ovary (CHO) cells, *E. coli*, and human embryonic kidney (293) cells. Both forms of TPO produced in these three hosts stimulated platelet production in mice, however, full length TPO isolated from CHO apeared to produce the greatest *in vivo* response. These results indicate that proper glycosylation of the carboxy-terminal domain may be necessary for optimal *in vivo* activity.

(a) *E. coli*-rhTPO(Met^{-1},153)

The "Met" form of the EPO domain (Met in the -1 position plus the first 153 residues of human TPO) produced in *E. coli* (see **Example 23**) was injected daily into

normal female C57 B6 mice as described in the legends to **Figs. 25A, 25B** and **25C**. These figures show that the non-glycosylated truncated form of TPO produced in *E. coli* and refolded as described above is capable of stimulating about a two-fold increase in platelet production in normal mice with out effecting the red or white blood cell population.

This same molecule injected daily into sublethally irradiated (^{137}Cs) female C57 B6 mice as described in the legends to **Figs. 26A, 26B** and **26C** stimulated platelet recovery and diminished nadir but had no effect on erythrocytes or leukocytes.

(b) CHO-rhTPO$_{332}$

The full length form of TPO produced in CHO and injected daily into normal female C57 B6 mice as described in the legends to **Figs. 27A, 27B** and **27C** produced about a five-fold increase in platelet production in normal mice with out effecting the erythrocyte or leukocyte population.

(c) CHO-rhTPO$_{332}$; *E. coli*-rhTPO$_{(Met^{-1},153)}$; 293-rhTPO$_{332}$; and *E. coli*-rhTPO$_{155}$

Dose response curves were constructed for treatment of normal mice with rhTPO from various cell lines (CHO-rhTPO$_{332}$; *E. coli*-rhTPO$_{(Met^{-1},153)}$; 293-rhTPO$_{332}$; and *E. coli*-rhTPO$_{155}$) as described in the legend to **Fig. 28**. This figure shows that all tested forms of the molecule stimulate platelet production, however the full length form produced in CHO has the greatest *in vivo* activity.

(d) CHO-rhTPO$_{153}$, CHO-rhTPO"clipped" and CHO-rhTPO$_{332}$

Dose response curves were also constructed for treatment of normal mice with various forms of rhTPO produced in CHO (CHO-rhTPO$_{153}$, CHO-rhTPO"clipped" and CHO-rhTPO$_{332}$) as described in the legend to **Fig. 29**. This figure shows that all tested CHO forms of the molecule stimulate platelet production,but that the full length 70 Kda form has the greatest *in vivo* activity.

16. General Recombinant Preparation of *mpl* Ligand and Variants

Preferably *mpl* ligand is prepared by standard recombinant procedures which involve production of the *mpl* ligand polypeptide by culturing cells transfected to express *mpl* ligand nucleic acid (typically by transforming the cells with an expression vector) and recovering the polypeptide from the cells. However, it is optionally envisioned that the *mpl* ligand may be produced by homologous recombination, or with recombinant production methods utilizing control elements introduced into cells already containing DNA encoding the *mpl* ligand. For example, a powerful promoter/enhancer element, a suppressor, or an exogenous transcription modulatory element may be inserted in the genome of the intended host cell in

proximity and orientation sufficient to influence the transcription of DNA encoding the desired *mpl* ligand polypeptide. The control element does not encode the *mpl* ligand, rather the DNA is indigenous to the host cell genome. One next screens for cells making the receptor polypeptide of this invention, or for increased or decreased levels of expression, as desired.

Thus, the invention contemplates a method for producing *mpl* ligand comprising inserting into the genome of a cell containing the *mpl* ligand nucleic acid molecule a transcription modulatory element in sufficient proximity and orientation to the nucleic acid molecule to influence transcription thereof, with an optional further step comprising culturing the cell containing the transcription modulatory element and the nucleic acid molecule. The invention also contemplates a host cell containing the indigenous *mpl* ligand nucleic acid molecule operably linked to exogenous control sequences recognized by the host cell.

A. Isolation of DNA Encoding mpl ligand Polypeptide

The DNA encoding *mpl* ligand polypeptide may be obtained from any cDNA library prepared from tissue believed to possess the *mpl* ligand mRNA and to express it at a detectable level. The *mpl* ligand gene may also be obtained from a genomic DNA library or by *in vitro* oligonucleotide synthesis from the complete nucleotide or amino acid sequence.

Libraries are screened with probes designed to identify the gene of interest or the protein encoded by it. For cDNA expression libraries, suitable probes include monoclonal or polyclonal antibodies that recognize and specifically bind to the *mpl* ligand. For cDNA libraries suitable probes include oligonucleotides of about 20-80 bases in length that encode known or suspected portions of the *mpl* ligand cDNA from the same or different species; and/or complementary or homologous cDNAs or fragments thereof that encode the same or a similar gene. Appropriate probes for screening genomic DNA libraries include, but are not limited to, oligonucleotides, cDNAs, or fragments thereof that encode the same or a similar gene, and/or homologous genomic DNAs or fragments thereof. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures as described in Chapters 10-12 of Sambrook *et al., supra.*

An alternative means to isolate the gene encoding *mpl* ligand is to use PCR methodology as described in section 14 of Sambrook *et al., supra.* This method requires the use of oligonucleotide probes that will hybridize to DNA encoding the *mpl* ligand. Strategies for selection of oligonucleotides are described below.

A preferred method of practicing this invention is to use carefully selected oligonucleotide sequences to screen cDNA libraries from various tissues, preferably human or porcine kidney (adult or fetal) or liver cell lines. For example, human fetal

liver cell line cDNA libraries are screened with the oligonucleotide probes. Alternatively, human genomic libraries may be screened with the oligonucleotide probes.

The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The actual nucleotide sequence(s) is usually designed based on regions of the *mpl* ligand which have the least codon redundancy. The oligonucleotides may be degenerate at one or more positions. The use of degenerate oligonucleotides is of particular importance where a library is screened from a species in which preferential codon usage is not known.

The oligonucleotide must be labeled such that it can be detected upon hybridization to DNA in the library being screened. The preferred method of labeling is to use ATP (*e.g.*, γ^{32}P) and polynucleotide kinase to radiolabel the 5' end of the oligonucleotide. However, other methods may be used to label the oligonucleotide, including, but not limited to, biotinylation or enzyme labeling.

Of particular interest is the *mpl* ligand nucleic acid that encodes a full-length *mpl* ligand polypeptide. In some preferred embodiments, the nucleic acid sequence includes the native *mpl* ligand signal sequence. Nucleic acid having all the protein coding sequence is obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence.

B. Amino Acid Sequence Variants of Native mpl ligand

Amino acid sequence variants of *mpl* ligand are prepared by introducing appropriate nucleotide changes into the *mpl* ligand DNA, or by *in vitro* synthesis of the desired *mpl* ligand polypeptide. Such variants include, for example, deletions from, or insertions or substitutions of, residues within the amino acid sequence for the porcine *mpl* ligand. For example, carboxy terminus portions of the mature full length *mpl* ligand may be removed by proteolytic cleavage, either *in vivo* or *in vitro*, or by cloning and expressing a fragment or the DNA encoding full length *mpl* ligand to produce a biologically active variant. Any combination of deletion, insertion, and substitution is made to arrive at the final construct, provided that the final construct possesses the desired biological activity. The amino acid changes also may alter post-translational processes of the *mpl* ligand, such as changing the number or position of glycosylation sites. For the design of amino acid sequence variants of the *mpl* ligand, the location of the mutation site and the nature of the mutation will depend on the *mpl* ligand characteristic(s) to be modified. The sites for mutation can be modified individually or in series, *e.g.*, by (1) substituting first with conservative amino acid choices and then with more radical selections depending upon the results achieved, (2) deleting the target residue, or (3) inserting residues of the same or a different class adjacent to the located site, or combinations of options 1-3.

A useful method for identification of certain residues or regions of the *mpl* ligand polypeptide that are preferred locations for mutagenesis is called "alanine scanning mutagenesis," as described by Cunningham and Wells, *Science,* 244:1081-1085 [1989]. Here, a residue or group of target residues are identified (*e.g.*, charged residues such as arg, asp, his, lys, and glu) and replaced by any, but preferably a neutral or negatively charged, amino acid (most preferably alanine or polyalanine) to affect the interaction of the amino acids with the surrounding aqueous environment in or outside the cell. Those domains demonstrating functional sensitivity to the substitutions then are refined by introducing further or other variants at or for the sites of substitution. Thus, while the site for introducing an amino acid sequence variation is predetermined, the nature of the mutation *per se* need not be predetermined. For example, to optimize the performance of a mutation at a given site, ala scanning or random mutagenesis is conducted at the target codon or region and the expressed *mpl* ligand variants are screened for the optimal combination of desired activity.

There are two principal variables in the construction of amino acid sequence variants: the location of the mutation site and the nature of the mutation. For example, variants of the *mpl* ligand polypeptide include variants from the *mpl* ligand sequence, and may represent naturally occurring alleles (which will not require manipulation of the *mpl* ligand DNA) or predetermined mutant forms made by mutating the DNA, either to arrive at an allele or a variant not found in nature. In general, the location and nature of the mutation chosen will depend upon the *mpl* ligand characteristic to be modified.

Amino acid sequence deletions generally range from about 1 to 30 residues, more preferably about 1 to 10 residues, and typically are contiguous. Alternatively, amino acid sequence deletions for the *mpl* ligand may include a portion of or the entire carboxy-terminus glycoprotein domain. Amino acid sequence deletions may also include one or more of the first 6 amino-terminus residues of the mature protein. Optional amino acid sequence deletions comprise one or more residues in one or more of the loop regions that exist between the 'helical bundels". Contiguous deletions ordinarily are made in even numbers of residues, but single or odd numbers of deletions are within the scope hereof. Deletions may be introduced into regions of low homology among the *mpl* ligands that share the most sequence identity to modify the activity of the *mpl* ligand. Or deletions may be introduced into regions of low homology among human *mpl* ligand and other mammalian *mpl* ligand polypeptides that share the most sequence identity to the human *mpl* ligand. Deletions from a mammalian *mpl* ligand polypeptide in areas of substantial homology with other mammalian *mpl* ligands will be more likely to modify the biological activity of the *mpl* ligand more

significantly. The number of consecutive deletions will be selected so as to preserve the tertiary structure of *mpl* ligands in the affected domain, *e.g.*, beta-pleated sheet or alpha helix.

Amino acid sequence insertions include amino- and/or carboxyl-terminal fusions ranging in length from one residue to polypeptides containing a hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions (*i.e.*, insertions within the mature *mpl* ligand sequence) may range generally from about 1 to 10 residues, more preferably 1 to 5, most preferably 1 to 3. An exemplary preferred fusion is that of *mpl* ligand or fragment thereof and another cytokine or fragment thereof. Examples of terminal insertions include mature *mpl* ligand with an N-terminal methionyl residue, an artifact of the direct expression of mature *mpl* ligand in recombinant cell culture, and fusion of a heterologous N-terminal signal sequence to the N-terminus of the mature *mpl* ligand molecule to facilitate the secretion of mature *mpl* ligand from recombinant hosts. Such signal sequences generally will be obtained from, and thus homologous to, the intended host cell species. Suitable sequences include STII or lpp for *E. coli*, alpha factor for yeast, and viral signals such as herpes gD for mammalian cells.

Other insertional variants of the *mpl* ligand molecule include the fusion to the N- or C-terminus of *mpl* ligand of immunogenic polypeptides (*i.e.*, not endogenous to the host to which the fusion is administered), *e.g.*, bacterial polypeptides such as beta-lactamase or an enzyme encoded by the *E. coli trp* locus, or yeast protein, and C-terminal fusions with proteins having a long half-life such as immunoglobulin constant regions (or other immunoglobulin regions), albumin, or ferritin, as described in WO 89/02922 published 6 April 1989.

A third group of variants are amino acid substitution variants. These variants have at least one amino acid residue in the *mpl* ligand molecule removed and a different residue inserted in its place. The sites of greatest interest for substitutional mutagenesis include sites identified as the active site(s) of *mpl* ligand and sites where the amino acids found in other analogues are substantially different in terms of side-chain bulk, charge, or hydrophobicity, but where there is also a high degree of sequence identity at the selected site among various *mpl* ligand species and/or within the various animal analogues of one *mpl* ligand member.

Other sites of interest are those in which particular residues of the *mpl* ligand obtained from various family members and/or animal species within one member are identical. These sites, especially those falling within a sequence of at least three other identically conserved sites, are substituted in a relatively conservative manner. Such conservative substitutions are shown in Table 3 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more

substantial changes, denominated exemplary substitutions in **Table 3**, or as further described below in reference to amino acid classes, are introduced and the products screened.

TABLE 3

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro	Pro
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe; norleucine	Leu
Leu (L)	norleucine; Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala	Leu
Pro (P)	Gly	Gly
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala; norleucine	Leu

Substantial modifications in function or immunological identity of the *mpl* ligand are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, Met, Ala, Val, Leu, Ile;

(2) neutral hydrophilic: Cys, Ser, Thr;

(3) acidic: Asp, Glu;

(4) basic: Asn, Gln, His, Lys, Arg;

(5) residues that influence chain orientation: Gly, Pro; and

(6) aromatic: Trp, Tyr, Phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

In one embodiment of the invention, it is desirable to inactivate one or more protease cleavage sites that are present in the molecule. These sites are identified by inspection of the encoded amino acid sequence, in the case of trypsin, *e.g.,* for an arginyl or lysinyl residue. When protease cleavage sites are identified, they are rendered inactive to proteolytic cleavage by substituting the targeted residue with another residue, preferably a basic residue such as glutamine or a hydrophobic residue such as serine; by deleting the residue; or by inserting a prolyl residue immediately after the residue.

In another embodiment, any methionyl residues other than the starting methionyl residue of the signal sequence, or any residue located within about three residues N- or C-terminal to each such methionyl residue, is substituted by another residue (preferably in accordance with **Table 3**) or deleted. Alternatively, about 1-3 residues are inserted adjacent to such sites.

Any cysteine residues not involved in maintaining the proper conformation of the *mpl* ligand also may be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant crosslinking. It has been found that the first and forth cysteines in the epo domain, numbered from the amino-terminus, are necessary for maintaining proper conformation but that the second and third are not. Accordingly, the second and third cysteines in the epo domain may be substituted.

Nucleic acid molecules encoding amino acid sequence variants of *mpl* ligand are prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of *mpl* ligand polypeptide.

Oligonucleotide-mediated mutagenesis is a preferred method for preparing substitution, deletion, and insertion variants of *mpl* ligand DNA. This technique is well known in the art as described by Adelman *et al.*, *DNA*, 2:183 [1983]. Briefly, *mpl*